



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 164310

TO: Jon E Angell
Location: 2d20 / 2c18
Friday, September 02, 2005
Art Unit: 1635
Phone: 571-272-0756
Serial Number: 09 / 972916

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

This Page Blank (uspto)

From: Chan, Christina
Sent: Tuesday, August 30, 2005 11:10 AM
To: Angell, Jon E; STIC-Biotech/ChemLib
Subject: RE: RUSH Sequence Search Request 09/972,916

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Angell, Jon E
Sent: Tuesday, August 30, 2005 11:24 AM
To: Chan, Christina
Subject: RUSH Sequence Search Request 09/972,916

I would like to request a RUSH sequence search because this case is an after final amended case that was forwarded to me 8/29/2005.

If approved, please forward the request to STIC and CC a copy to me, THANKS!

SEARCH REQUEST FORM

Scientific and Technical Information Center

Examiner# : 78697
Art Unit : 1635
Phone Number: 571-272-0756
Date: 8/30/2005
Serial Number: 09/972,916 (THULE et al.)
Mailbox & Bldg/Room Location: REMSEN 2C18
Results Format Preferred (circle): Paper

I would like to have a **standard** search of the nucleic acid databases including the pending applications databases (i.e., an **interference** search) performed using the following SEQ. ID NOs. from application : 09/972,916

SEQ ID NO: 1 (nucleic acid 51 nucleotides in length)
SEQ ID NO: 2 (nucleic acid 219 nucleotides in length)
SEQ ID NO: 3 (nucleic acid 270 nucleotides in length)
SEQ ID NO: 4 (nucleic acid 321 nucleotides in length)
SEQ ID NO: 5 (nucleic acid 372 nucleotides in length)
SEQ ID NO: 6 (nucleic acid 423 nucleotides in length)

Please note that the above sequence are gene expression regulatory sequences and the claims encompass nucleic acid sequences comprising any 1 or more of the above sequences.

STAFF USE ONLY

Searcher: Q
Searcher Phone: 2- 2504
Date Searcher Picked up: 8/30/05
Date Completed: 9/2/05
Searcher Prep/Rev. Time: 15
Online Time: 15

Type of Search

NA#: ✓ AA#:
Interference: ✓ SPDI:
S/L: Oligomer:
Encode/Transl:
Structure#: Text:
Inventor: Litigation:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIS:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

Please contact me by phone or email if you have any questions.

Thanks,
Eric

J. Eric Angell
Art Unit 1635
Office: REMSEN 2D20
mailbox: REM 2C18
571-272-0756

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 12:52:03 ; Search time 232.225 Seconds
(without alignment)
5582.626 Million cell updates/sec

Title: US-09-972-916b-2

Perfect score: 219
Sequence: 1 tcacagcgaacaaacta.....ttccgctactagctagccgc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20028:*
6: geneeqn20038:*
7: geneeqn20048:*
8: geneeqn20058:*
9: geneeqn20068:*
10: geneeqn20078:*
11: geneeqn20088:*
12: geneeqn20098:*
13: geneeqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219	100.0	219	6	ABX15372 Rat insul
2	219	100.0	270	6	ABX15373 Rat insul
3	219	100.0	321	6	ABX15374 Rat insul
4	219	100.0	372	6	ABX15375 Rat insul
5	219	100.0	423	6	ABX15376 Rat insul
6	219	100.0	474	6	ABX15377 Rat insul
7	219	100.0	525	6	ABX15378 Rat insul
8	219	100.0	576	6	ABX15379 Rat insul
9	219	100.0	627	6	ABX15380 Rat insul
10	219	100.0	678	6	ABX15381 Rat insul
11	219	100.0	729	6	ABX15382 Rat insul
12	219	100.0	780	6	ABX15383 Rat insul
13	219	100.0	831	6	ABX15384 Rat insul
14	219	100.0	882	6	ABX15385 Rat insul
15	219	100.0	933	6	ABX15386 Rat insul
16	219	100.0	984	6	ABX15387 Rat insul
17	219	100.0	1035	6	ABX15388 Rat insul
18	219	100.0	1086	6	ABX15389 Rat insul
19	219	100.0	1137	6	ABX15390 Rat insul
20	219	100.0	1188	6	ABX15391 Rat insul

ALIGNMENTS

RESULT 1	ABX15372	standard; DNA, 219 BP.
XX	ABX15372;	
XX	17-APR-2003 (first entry)	
XX	Rat insulin-sensitive element (ISE) DNA.	
DE	Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGF1P-1;	
XX	insulin-sensitive element; ISE; basal promoter; hypoglycaemia; insulin;	
KW	insulin-like growth factor binding protein-1; hypoglycaemia; glucose;	
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;	
KW	glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoadicidosis;	
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;	
KW	intracellular hormone receptor; insulin regulator construct; anabolic.	
XX	Rattus norvegicus.	
OS	US2002107198-A1.	
XX	08-AUG-2002.	
PD	10-OCT-2001; 2001US-00972916.	
PF	11-OCT-2000; 2000US-0239113P.	
PR	(THUL/) THUL P M.	
PA	Thule PM;	
XX	WPI; 2002-674190/72.	
DR	New insulin regulator cassette, useful e.g. for treating diabetes,	
PT	provides specific, glucose-inducible transgenic expression of insulin in	
PT	liver cells.	
PS	Claim 8; Page 13; 37pp; English.	
XX	The invention relates to an insulin regulator construct comprising a	
CC	glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene	
CC	promoter and an insulin-sensitive element (ISE) of an insulin-like growth	

CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 100.0%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.5e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
DB 1 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
QY 61 TTAAACCCGCTGCGGAGCCAGCCCTTCTAAGGCCCTGGATGAGCCGACATGAT 120
DB 61 TTAAACCCGCTGCGGAGCCAGCCCTTCTAAGGCCCTGGATGAGCCGACATGAT 120
QY 121 CCACCTGCGCCGCGAGACACAAACCCAGCAGCATTTGAACATGACACGGCCATGCGCC 180
DB 121 CCACCTGCGCCGCGAGACACAAACCCAGCAGCATTTGAACATGACACGGCCATGCGCC 180
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219

RESULT 2

ABX15373
ID ABX15373 standard; DNA; 270 BP.

AC ABX15373;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #1.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
OS Synthetic.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THUL P M.

XX Thule PM;

XX

DR WPI: 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

XX Claim 9; Page 13-14; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention

SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 100.0%; Score 219; DB 6; Length 270;
Best Local Similarity 100.0%; Pred. No. 2.7e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
DB 52 TCACAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 111
QY 61 TTAAACCCGCTGCGGAGCCAGCCCTTCTAAGGCCCTGGATGAGCCGACATGAT 120
DB 112 TTAAACCCGCTGCGGAGCCAGCCCTTCTAAGGCCCTGGATGAGCCGACATGAT 171
QY 121 CCACCTGCGCCGCGAGACACAAACCCAGCAGCATTTGAACATGACACGGCCATGCGCC 180
DB 172 CCACCTGCGCCGCGAGACACAAACCCAGCAGCATTTGAACATGACACGGCCATGCGCC 231
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219
DB 232 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270

RESULT 3

ABX15374
ID ABX15374 standard; DNA; 321 BP.

AC ABX15374;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #2.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
OS Synthetic.

PN US2002107198-A1.

XX 08-AUG-2002.
 PD 10-OCT-2001; 2001US-00972916.
 XX 11-OCT-2000; 2000US-0239113P.
 XX (THUL/) THULE P M.
 PA Thule PM;
 PI WPI; 2002-674190/72.
 XX New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX Claim 9; Page 14; 37pp; English.
 XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 XX Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 219; DB 6; Length 321;
 Best Local Similarity 100.0%; Pred. No. 2.9e-60;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTAGACGCTGCTGACAAATCA 60
 DB 103 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTAGACGCTGCTGACAAATCA 162
 QY 61 TTAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 120
 DB 163 TTAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 222
 QY 121 CCACTGCCGCGGAGACACAAACCCAGCGAGTTGAACACTGACACAGGCCATCTGGCCC 180
 DB 223 CCACTGCCGCGGAGACACAAACCCAGCGAGTTGAACACTGACACAGGCCATCTGGCCC 282
 QY 181 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 219
 DB 283 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 321
 RESULT 4
 ABX15375
 ID ABX15375 standard; DNA; 372 BP.
 XX
 AC ABX15375;
 XX
 DT 17-APR-2003 (first entry)
 XX
 DE Rat insulin regulator construct DNA #3.
 XX
 KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;

KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX US2002107198-A1.
 XX 08-AUG-2002.
 PD 10-OCT-2001; 2001US-00972916.
 XX 11-OCT-2000; 2000US-0239113P.
 XX (THUL/) THULE P M.
 PA Thule PM;
 PI WPI; 2002-674190/72.
 XX New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX Claim 9; Page 14; 37pp; English.
 XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 XX Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 219; DB 6; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-60;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTAGACGCTGCTGACAAATCA 60
 DB 154 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTAGACGCTGCTGACAAATCA 213
 QY 61 TTAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 120
 DB 214 TTAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 273
 QY 121 CCACTGCCGCGGAGACACAAACCCAGCGAGTTGAACACTGACACAGGCCATCTGGCCC 180
 DB 274 CCACTGCCGCGGAGACACAAACCCAGCGAGTTGAACACTGACACAGGCCATCTGGCCC 333
 QY 181 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 219
 DB 334 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 372

RESULT 5
 ABX15376
 ID ABX15376 standard; DNA; 423 BP.
 AC
 XX ABX15376;
 XX
 DT 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #4.
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KM insulin-sensitive element; ISE; basal promoter; hypoglycaemia; insulin;
 KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX
 PN US2002107198-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR MPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
 Query Match 100.0%; Score 219; DB 6; Length 423;
 Best Local Similarity 100.0%; Pred.No. 3,1e-60;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCACAGCAAAACAACTATTATTTGAACGCGGATCCAGACCGCTGCGCATCA 60
 DB 205 TCACAGCAAAACAACTATTATTTGAACGCGGATCCAGACCGCTGCGCATCA 264
 61 TTAACCGTGCTGCGAGCGACCGCTTCATAAGCGCTGGGTATGCGCAGCCAGCATGAT 120

DB 265 TTAACCGTGCTGCGAGCGACCGCTTCATAAGCGCTGGGTATGCGCAGCCAGCATGAT 324
 QY 121 CCACTGCCCGCGGAGACAAACCCAGAGCATTTGAACACTGCACAGGCATGTGCC 180
 DB 325 CCACTGCCCGCGGAGACAAACCCAGAGCATTTGAACACTGCACAGGCATGTGCC 384
 QY 181 AGAGAGCTGTGACACACACTTCGCTACTAGTACCGC 219
 DB 365 AGAGAGCTGTGACACACACTTCGCTACTAGTACCGC 423
 RESULT 6
 ABK63701
 ID ABK63701 standard; cDNA; 1500 BP.
 AC
 XX ABK63701;
 XX
 DT 18-JUN-2002 (first entry)
 DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
 XX
 KM Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KM differential expression; centrilobular necrosis; steatosis.
 XX
 OS Rattus norvegicus.
 OS
 PN WO200210453-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-US023872.
 XX
 PR 31-JUL-2000; 2000US-0222040P.
 PR 02-NOV-2000; 2000US-0244888P.
 PR 11-MAY-2001; 2001US-0230029P.
 PR 15-MAY-2001; 2001US-0230645P.
 PR 22-MAY-2001; 2001US-0232336P.
 PR 06-JUN-2001; 2001US-0235798P.
 PR 13-JUN-2001; 2001US-0297457P.
 PR 19-JUN-2001; 2001US-0238888P.
 PR 09-JUL-2001; 2001US-030459P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
 XX
 DR MPI; 2002-241625/29.
 XX
 CC Predicting toxic effects of compounds or the progression of these toxic
 CC effects by determining the changes in gene expression in tissues or cells
 CC exposed to the toxin and comparing these to gene expression in unexposed
 CC tissues or cells.
 XX
 PS Claim 1; SEQ ID NO 1608; 239pp; English.
 XX
 CC The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic effect
 CC of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression. The
 CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilise a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,

CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterized by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX

SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 137 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTACAC 67

OY 197 CACTTCGCGCTACTAGCTA 214
DB 68 CACTTCGCGCTACTATCTA 85

RESULT 7
ADB58201
ID ADB58201 standard; DNA; 1500 BP.

AC ADB58201;

DT 04-DEC-2003 (first entry)

DE Toxicity-related gene, SEQ ID 3227.

KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;

KW drug screening; toxicity assay; ds.

OS Unidentified.

XX WO2003064624-A2.

PD 07-AUG-2003.

PF 31-JAN-2003; 2003WO-US003194.

PR 31-JAN-2002; 2002US-00060087.

PR 15-MAR-2002; 2002US-0364045P.

PR 15-MAR-2002; 2002US-0364055P.

PR 30-DEC-2002; 2002US-0436643P.

XX (GENE-) GENE LOGIC INC.

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX WPI; 2003-689530/65.

PT Predicting a toxic effect of a compound, useful in identifying toxicity

XX markers in liver tissues or cells for drug screening and toxicity assays,

XX comprises preparing gene expression profile of tissue or cells exposed to

XX the compound.

XX Claim 1, SEQ ID NO 3227; 1156bp; English.

XX The present invention relates to a method for predicting a toxic effect

XX of a compound. The method comprises preparing a gene expression profile

XX of a tissue or cell sample exposed to the compound, and comparing the

XX gene expression profile to a database comprising SEQ ID 1-4925, where

XX differential expression of the gene indicates at least one toxic effect.

CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 34.9%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 137 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTACAC 67

OY 197 CACTTCGCGCTACTAGCTA 214
DB 68 CACTTCGCGCTACTATCTA 85

RESULT 8
ADB52710
ID ADB52710 standard; DNA; 1500 BP.

AC ADB52710;

DT 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.

KM toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

KW toxicity marker; toxicity progression; drug screening;

XX primary rat hepatocyte toxicity modelling; gene; ds.

OS Rattus norvegicus.

XX WO2003065993-A2.

PD 14-AUG-2003.

PF 04-FEB-2003; 2003WO-US003482.

PR 04-FEB-2002; 2002US-0353171P.

PR 13-MAR-2002; 2002US-0363534P.

PR 08-APR-2002; 2002US-0370248P.

PR 10-APR-2002; 2002US-0371134P.

PR 10-APR-2002; 2002US-0371135P.

PR 11-APR-2002; 2002US-0371413P.

PR 19-APR-2002; 2002US-0373601P.

PR 19-APR-2002; 2002US-0373602P.

PR 22-APR-2002; 2002US-0374139P.

PR 08-MAY-2002; 2002US-0378652P.

PR 09-MAY-2002; 2002US-0378653P.

PR 09-MAY-2002; 2002US-0378655P.

PR 09-JUL-2002; 2002US-0394230P.

PR 09-JUL-2002; 2002US-0394253P.

PR 04-SEP-2002; 2002US-0407688P.

PR 28-JAN-2003; 2003US-0442900P.

XX (GENE-) GENE LOGIC INC.

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;

XX Elashoff M;

DR WPI; 2003-731472/69.

XX Determining if a compound induces a toxic effect on a tissue or cell, for

XX identifying hepatotoxic compounds, comprises comparing a gene expression

PT profile of a tissue or cell sample to a database of tox mean and non-tox

PT mean values.

XX Claim 44; SEQ ID NO 3252; 874bp; English.

XX The present invention describes a method for determining whether a

CC compound induces a toxic effect on a tissue or cell. The method comprises

CC preparing a gene expression profile of a tissue or cell sample exposed to

CC the compound, and comparing the gene expression profile to a database

CC comprising data or information on the Tox mean and non-Tox mean value.

CC The method is useful for predicting or identifying at least one toxic

CC effect, particularly hepatotoxicity, of a test or unknown compound. The

CC genes listed in the specification are useful as diagnostic or toxicity

CC markers for the prediction or identification of the physiological state

CC of tissue or cell sample that has been exposed to a compound, or to

CC identify or predict the toxic effects of a compound or an agent. These

CC may also be used as markers for monitoring toxicity progression or for

CC drug screening. The present sequence represents a primary rat hepatocyte

CC toxicity modelling related gene sequence from the present invention.

SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

XX

XX Query Match 34.9%; Score 76.4; DB 10; Length 1500;

XX Best Local Similarity 98.7%; Pred. No. 3.6e-14;

XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGATTGAACACTGACAGGCGCATCTGCCCAGAGAGCTGTGACAC 196

DB 8 CACAAACCCAGGAGATTGAACACTGACAGGCGCATCTGCCCAGAGAGCTGTGACAC 67

QY 197 CACTTCCGCTACTAGCTA 214

DB 68 CACTTCCGCTACTAGCTA 85

XX

XX RESULT 9

XX ABT41911

XX ID ABT41911 standard; DNA; 1500 BP.

XX

XX ABT41911;

XX

XX 26-JUN-2003 (first entry)

XX

XX Toxicity modelling related rat gene SEQ ID No 1613.

XX

XX Toxic effect; gene expression profile; renal toxicity; toxicity marker;

XX database; drug screening; toxicity assay; rat; ds.

XX

XX Rattus norvegicus.

XX

XX WO200295000-A2.

XX

XX 28-NOV-2002.

XX

XX 22-MAY-2002; 2002WO-US016173.

XX

XX 22-MAY-2001; 2001US-0292335P.

XX

XX 13-JUN-2001; 2001US-0297523P.

XX

XX 19-JUN-2001; 2001US-0298925P.

XX

XX 10-JUL-2001; 2001US-0303807P.

XX

XX 10-JUL-2001; 2001US-0303808P.

XX

XX 10-JUL-2001; 2001US-0303810P.

XX

XX 28-AUG-2001; 2001US-0315047P.

XX

XX 27-SEP-2001; 2001US-0324928P.

XX

XX 22-OCT-2001; 2001US-0330462P.

XX

XX 01-NOV-2001; 2001US-0330867P.

XX

XX 21-NOV-2001; 2001US-0331805P.

XX

XX 06-DEC-2001; 2001US-0336144P.

XX

XX 19-DEC-2001; 2001US-0340873P.

PR 21-FEB-2002; 2002US-0357842P.

PR 21-FEB-2002; 2002US-0357843P.

PR 21-FEB-2002; 2002US-0357844P.

PR 15-MAR-2002; 2002US-0364134P.

PR 08-APR-2002; 2002US-0370144P.

PR 08-APR-2002; 2002US-0370206P.

PR 08-APR-2002; 2002US-0370247P.

PR 17-APR-2002; 2002US-0372949P.

PR 21-APR-2002; 2002US-0371679P.

XX

XX (GENE-) GENE LOGIC INC.

XX

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX WPI; 2003-148464/14.

XX

XX Predicting at least one toxic effect of a compound, useful for toxicity

PT modeling, comprises preparing a gene expression profile of a tissue or

PT cell sample exposed to the compound, and comparing the gene expression

PT profile to a database.

XX

XX Example 4; Page; 446pp; English.

XX

XX The invention relates to a novel method of predicting at least one toxic

CC effect of a compound. The method comprises a gene expression profile of a

CC tissue or cell sample exposed to the compound, and comparing the gene

CC expression profile to a database comprising at least part of the data or

CC information given in the specification. The methods are useful for

CC predicting at least one toxic effect of a compound, predicting the renal

CC progression of a toxic effect of a compound, predicting the renal

CC toxicity of a compound, or identifying toxicity markers in tissues or

CC cells exposed to known renal toxin. The genes are useful as toxicity

CC markers in drug screening and toxicity assays, in monitoring disease or

CC physiological states, or disease progression. This polynucleotide

CC represents a rat DNA sequence relating to the toxic effect database

CC described in the specification. NOTE: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the World Intellectual Property

CC Organization

XX

XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

SQ

XX

XX Query Match 34.9%; Score 76.4; DB 10; Length 1500;

XX Best Local Similarity 98.7%; Pred. No. 3.6e-14;

XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGATTGAACACTGACAGGCGCATCTGCCCAGAGAGCTGTGACAC 196

DB 8 CACAAACCCAGGAGATTGAACACTGACAGGCGCATCTGCCCAGAGAGCTGTGACAC 67

QY 197 CACTTCCGCTACTAGCTA 214

DB 68 CACTTCCGCTACTAGCTA 85

XX

XX RESULT 10

XX ADP72689

XX ID ADP72689 standard; DNA; 1500 BP.

XX

XX ADP72689;

XX

XX 26-AUG-2004 (first entry)

XX

XX Renal toxin progression gene marker #1278.

XX

XX ds; toxic effect; gene expression profile; kidney tissue;

XX differential gene expression; toxicity progression; toxicity marker;

XX drug screening; toxicity assay; kidney pathology; nephritis;

XX kidney necrosis; glomerular injury; tubular injury;

XX focal segmental glomerulosclerosis.

XX

XX Rattus norvegicus.

XX

PN WO2004048598-A2.
XX
PD 10-JUN-2004.
XX
PF 24-NOV-2003; 2003WO-US037556.
XX
PR 22-NOV-2002; 2002US-00301856.
XX
PA (GENE-) GENE LOGIC INC.
PI Mendrick DL, Porter MM, Johnson KR, Castle A, Higge B;
PI Elashoff M;
XX
DR WPI; 2004-460771/43.
XX
PT Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
PS Claim 11: SEQ ID NO 1278; 266pp; English.
XX
CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
Query Match 34.9%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 137 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCGCAGAGAGCTGTGACCAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCGCAGAGAGCTGTGACCAC 67
XX
QY 197 CACTTCCGCTACTAGCTA 214
DB 68 CACTTCCGCTACTAGCTA 85
XX
RESULT 11
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
AC AAH22429;
XX
DT 22-AUG-2001 (first entry)
XX
DE Rat insulin-like growth factor binding protein nucleotide sequence.
XX
KW Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW non-steroidal antiinflammatory drug; ds.

XX
OS Rattus norvegicus.
XX
PN WO200138579-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032049.
XX
PR 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gould-Rothberg BE, Dipippo VA, Rambeh TM, Gerwein RW;
XX
DR WPI; 2001-355948/37.
XX
XX
PT Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
PS Disclosure: Page 22-24; 76pp; English.
XX
CC The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in a reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;
XX
Query Match 34.9%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 5.3e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 137 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCGCAGAGAGCTGTGACCAC 196
DB 1 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCGCAGAGAGCTGTGACCAC 60
XX
QY 197 CACTTCCGCTACTAGCTA 214
DB 61 CACTTCCGCTACTAGCTA 78
XX
RESULT 12
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
AC ABN95896;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2394 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.

```

PN WO200229103-A2.
XX
XX 11-APR-2002.
PD
XX 02-OCT-2001; 2001WO-US030589.
PF
XX 02-OCT-2000; 2000US-0237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG,
XX WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer, hepatocellular
XX carcinoma or metastatic liver tumor in a patient, involves detecting the
XX level of expression of two or more genes in a liver tissue sample.
XX
XX Claim 1; SEQ ID NO 2394; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
SQ
Query Match 27.8%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 5.9e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
QY 2 CACAAGCAAAACAACTATTATTTGAACACGGGAGTCCTAGACGCTGCCTGACAAATCAT 61
DB 473 CACTAGCAAAACAACTATTATTTGAACACTCAGCTCTAGCGTGCGGCGCTGCCAATCAT 532
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGATGCGCAGC 111
DB 533 TAACCTCTGTGTGCAAGTGCGCGGCTGTGCTTTTAAAGTGCGCGCTGTGTCAGC 592
QY 112 CAGCATGGTCCAGTGCCTGCGCGGAGACACAAACCGGCGAGCTTGAACACTGSC-ACACGG 170
DB 593 GAGCATTCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 644
QY 171 CCATCTGCCGAGAGAGCTGTGACCAACCACTTC 202
DB 645 CCACCTCCCGAGAGAGCACTGCGCACCGCTCC 676
RESULT 13
ABV75371
ID ABV75371 standard; DNA; 6128 BP.
XX
XX ABV75371;
XX
XX 07-MAR-2003 (first entry)
XX
XX Human IGFBP-1 gene sequence.
XX
XX Insulin-like growth factor binding protein; IGFBP; cytotrophic; liver;
XX cancer; human; IGFBP-1; gene; ds.
XX
XX Homo sapiens.
OS

```

```

XX Key Location/Qualifiers
FH 762..5200
FT CDS
FT
FT /tag= a
FT /product= "IGFBP-1"
FT /note= "insulin-like growth factor binding protein;
FT contains introns"
FT 762..1110
FT /tag= b
FT /number= 1
FT intron 1111..2656
FT /tag= c
FT /number= 1
FT exon 2657..2826
FT /tag= d
FT /number= 2
FT intron 2827..4040
FT /tag= e
FT /number= 2
FT exon 4041..4169
FT /tag= f
FT /number= 3
FT intron 4170..5068
FT /tag= g
FT /number= 3
FT exon 5069..5197
FT /tag= h
FT /number= 4
XX
XX WO200290580-A1.
XX
XX 14-NOV-2002.
XX
XX 03-MAY-2002; 2002WO-AU000558.
XX
XX 03-MAY-2001; 2001US-0288441P.
XX
XX (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
XX (SICE-) SINGAPORE GEN HOSPITAL PTE LTD.
XX (ARGA/) ARGAE V P.
XX
XX Huynh TH, Chow PKH, Soo KC;
XX WPI; 2003-103522/09.
XX P-PSDB; ABB82757.
XX
XX Detecting the presence or diagnosing the risk of a liver cancer in a
XX patient comprises detecting aberrant expression of a gene encoding an
XX insulin-like growth factor binding protein.
XX
XX Example; Page 104-108; 142pp; English.
XX
XX The invention relates to detecting the presence or diagnosing the risk of
XX a liver cancer in a patient. The method involves detecting in a
XX biological sample obtained from the patient aberrant expression of a gene
XX encoding an insulin-like growth factor binding protein (IGFBP). The
XX method is useful for detecting the presence or diagnosing the risk of a
XX liver cancer or for screening agents in a patient. The agent is useful
XX for the manufacture of a medicament for treating and/or preventing liver
XX cancer. The present sequence represents a human IGFBP-1 polypeptide
XX encoding genomic DNA (GenBank Accession No. M74587)
XX
XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
SQ
Query Match 27.8%; Score 60.8; DB 10; Length 6128;
Best Local Similarity 66.5%; Pred. No. 5.9e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
QY 2 CACAAGCAAAACAACTATTATTTGAACACGGGAGTCCTAGACGCTGCCGCAATCAT 61
DB 473 CACTAGCAAAACAACTATTATTTGAACACTCAGCTCTAGCGTGCGGCGCTGCCAATCAT 532
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGATGCGCAGC 111

```

D _b	533	TAACTCCTCGTGCAGATGACGGCGGTGCTCCCTTTATAAGTGCGCGCGCTTCACGC	592
Q _y	112	CAGCATGTCCTACTGCCCCGCCGAGACAACCCAGCAGATTGAACATG-C-ACACGG	170
D _b	593	GAGCATGGCGCAACGGCATTC-----CATTCAGCGAGAGTCTGCGCCGCCGCCCG	644
Q _y	171	CCATCTGCCAAGAGCTGTGACCACCACTTC	202
D _b	645	CCAACCTCCCAGAGACACTGGCGCACCGCTCC	676

RESULT 14
AAH57489
ID AAH57489 standard; cDNA; 6134 BP.

AC	AAH57489;	
XX		
DT	10-SEP-2001	(first entry)
XX		
BE	Human liver cell specific cDNA sequence	SEQ ID NO:329

KM Human, tissue specific; diagnosis; brain; heart; skeletal muscle; lung
 KM liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KM metabolic disease; developmental disease; cytostatic; immunomodulatory
 KM neuroprotective; gene therapy; cancer; immunopathology; neuropathology

OS Homo sapiens.

PN WO200132927-A2.

PD 10-MAY-2001

02-NOV-2000; 2000WO-US030396

04-NOV-1999: 99US-0163508P

PA (INCY-) INCYTE GENOMICS INC.

XX Sornaae T. Selfhamer JT Watson GA.
PT

XX
DB WPT: 2001-291057/30

[illegible]

PT prognosis or monitoring of treatments where the gene is associated with a disease.

XX
2

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (1). (1) can have cytoskeletal, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (1) and proteins (II) encoded by them are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (1) in a sample indicates the differentiation of embryonic stem cells into a tissue selected from brain, heart, kidney, liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used to produce an expression profile that defines a metabolic or developmental process, treatment, condition, disease or disorder. The gene profile can be used for diagnosis, prognosis or monitoring of treatments and for investigating a predisposition to a disorder where the gene is associated with a cancer, immunopathology or neuropathology

Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;

Query Match 27.8%; Score 60.8; DB 4; Length 6134;

```

      00.58;  File: NO. 3.26-05;
Matches 141;  Conservative  0;  Mismatches  52;  Indels  19;  Gaps  3;

```

2 CACAAGCAAACTTATTTTGACACACGGGATCCTAGCACCCTGCCCTGACAATCAT 61

27

Db	473	CAC	TAG	CAAAA	CA	AACTT	ATTTT	TG	AAC	ACT	AGCT	CCT	AG	CGT	CGG	CGT	CGCA	TAT	532				
Oy	62	TAA	CCC	-----	-----	GTG	CTG	CCG	AG	CG	AGC	CGCTT	CTA	TAG	CGC	CTG	GGT	ATG	CGC	AGC	111		
Db	533	TAA	CTC	CTG	TGC	CA	AGT	GG	CG	CGC	CT	GTG	CGCTT	TAA	GTG	CGC	CT	GT	CTC	AGC	532		
Oy	112	CAG	CAT	GTG	TC	ACT	GC	CGC	CG	AG	CA	CA	AA	CC	AG	CAG	ATT	GA	CA	TGC	-ACA	CGG	170
Db	593	GAG	CAT	GGG	CA	CCG	CG	CA	TCC	-----	-----	CAT	CC	AG	CAG	CAT	CTG	CGC	CGC	CGC	CGC	CGG	644
Oy	171	CC	AT	TCC	CA	AG	AG	CT	GT	GA	CA	CA	CA	CTT	202								
Db	645	CCA	CC	CTT	CCA	GA	GA	CA	CT	GG	CA	CA	CC	CTC	676								

ABL32997/c
ID ABL32997 standard; DNA; 7061 BP.

AC ABL32997;

DT 26-MAR-2002 (first entry)

Human immune system associated gene SBO ID NO: 970.

Human: immune disease; cytosine methylation; antiaesthetic;
antiarthrogenic; antianemic; cytosine; neotropic;
neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
antihemetic; antiarthritic; antidiabetic; antipsychotic;
antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
gene;

XX 3

OS Homo sapiens.

PN WO200200928-A2

03-JAN-2002

02-JUL-2001: 2001WO-EP007537.

30-JUN-2000: 2000DE-01032529

PR 01-SEP-2000; 2000DE-01043826.
XX

PA (EPiG-) EPIGENOMICS AG.
XX

PI Olek A, Piepenbrock C, Berlin K;
yy

WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for

PT methylation.

PS Claim 1; SEQ ID NO 970; 32pp + Sequence Listing; German.

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention

Sequence 7061 BP; 1852 A; 147 C; 1865 G; 3197 T; 0 U; 0 Other;

Query Match 17.38; Score 37.8; DB 6; Length 7061;

```

      100% Identity (15.00) 1200: NO. 0.10;
      100% Identity (15.00) 1200: NO. 0.10;
Matches 48; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

2 CACAAGCAAACTTATTTTGAACACGGGATCCTAGCAGCGTGCCCTGACAATCAT 61

Db	2185	CACTAACAACTTAATTTAACTCAACTCTACTACGACGACGCTACCAATCAT	2126
Qy	62	TAACC	66
Db	2125	TAACC	2121

Search completed: September 1, 2005, 00:34:41
Job time : 238.225 secs

Query Match	100.0%;	Score 51;	DB 3;	Length 194;
-------------	---------	-----------	-------	-------------

Best Local Similarity 100.0%; Pred. No. 6.3e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 11 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 61

RESULT 2

US-08-791-849A-14
; Sequence 14, Application US/08791849A
; Patent No. 5914449
; GENERAL INFORMATION:
; APPLICANT: Makoto MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,849A
; FILING DATE: January 30, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: rat (Rattus norvegicus)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(3212..3218, 3766..3948, 5917..6008,
; 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; LOCATION: 9298..9479, 10163..10269)
; FEATURE:
; NAME/KEY: exon
; LOCATION: join(3194..3218, 3766..3948, 5917..6008,
; 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; FEATURE:
; NAME/KEY: intron
; LOCATION: join(3219..3765, 3949..5916, 6009..6151,
; 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
; LOCATION: 9480..10162)
US-08-791-849A-14

Query Match 100.0%; Score 51; DB 2; Length 13011;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||

DB 3021 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 3071

RESULT 3

US-09-949-016-11786
; Sequence 11786, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11786
; LENGTH: 15108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11786

Query Match 71.8%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.0012;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 2357 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 2399

RESULT 4

US-09-949-016-17205
; Sequence 17205, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 17205
; LENGTH: 15108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17205

Query Match 71.8%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.0012;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 2357 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 2399

RESULT 5
US-09-234-332-5
Sequence 5, Application US/09234332A
Patent No. 6087168
GENERAL INFORMATION:
APPLICANT: Cedars-Sinai Medical Center
APPLICANT: Michael F. Levesque, M.D.
APPLICANT: Thomas Neuman, Ph.D.
TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
TITLE OF INVENTION: NEURONS; TRANSDIFFERENTIATION OF EPIDERMAL CELLS
FILE REFERENCE: P07 41494
CURRENT APPLICATION NUMBER: US/09/234,332A
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 3138
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: gene
LOCATION: (0)...(0)
OTHER INFORMATION: zfc 1 Protein gene; Genbank Accession D76435
US-09-234-332-5

Query Match 47.8%; Score 24.4; DB 3; Length 3138;
Best Local Similarity 73.8%; Pred. No. 20;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 GCGACGGGACATCCCGTGGTCTCGGACTCGGCCCCCAGT 48
DB 764 GCGCCGACGACGACGATCTCTGACGCGGCCCCCAGT 805

RESULT 6
US-08-945-140-2
Sequence 2, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
US-08-945-140-2

Query Match 47.1%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GCGACGGGACATCCCGTGGTTC 29
DB 1 GCGACGGGACATCCCGTGGTTC 24

RESULT 7
US-08-775-428-1/c
Sequence 1, Application US/08775428
Patent No. 5976834
GENERAL INFORMATION:
APPLICANT: Sathe, Ganesh
APPLICANT: Fueterer, Wendy
APPLICANT: Bergema, Derek
APPLICANT: Ellis, Catharine
TITLE OF INVENTION: CDNA CLONE HNF1D15 THAT ENCODES
TITLE OF INVENTION: A NOVEL HUMAN 7-TRANSMEMBRANE RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,428
FILING DATE: 09-JAN-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50042
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4060
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1498 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-775-428-1

Query Match 47.1%; Score 24; DB 2; Length 1498;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1956
LENGTH: 3078
TYPE: DNA
ORGANISM: Human
US-09-949-016-1956

Query Match 45.9%; Score 23.4; DB 4; Length 3078;
Best Local Similarity 73.2%; Pred. No. 44;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 11 CGGGGCACTCCGCTGCTCTGACTGCGCCCGCAGTGA 51
DB 501 CGAGGCACTGACGAGAGCGCTGAGTGTGCGCCCGCAGTGA 461

RESULT 12
US-09-023-655-1294/c
Sequence 1294, Application US/09023655
Patent No. 6607879

GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1294:
SEQUENCE CHARACTERISTICS:
LENGTH: 3088 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 931667
US-09-023-655-1294

Query Match 45.9%; Score 23.4; DB 4; Length 3088;
Best Local Similarity 73.2%; Pred. No. 44;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 11 CGGGGCACTCCGCTGCTCTGACTGCGCCCGCAGTGA 51
DB 501 CGAGGCACTGACGAGAGCGCTGAGTGTGCGCCCGCAGTGA 461

DB 502 CGAGGCACTGACGAGAGCGCTGAGTGTGCGCCCGCAGTGA 462

RESULT 13
US-09-949-016-13508/c
Sequence 13508, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13508
LENGTH: 46244
TYPE: DNA
ORGANISM: Human
US-09-949-016-13508

Query Match 45.9%; Score 23.4; DB 4; Length 46244;
Best Local Similarity 67.3%; Pred. No. 59;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 CATGGGCGACGAGGCGCACTCCGCTGCTCTGACTGCGCCCGCAGTGA 49
DB 39370 CTTGGGCTCAGGGGCGCACTCTCGGCTGCTCTGCTGCGCCCGCTGG 39322

RESULT 14
US-09-949-016-49730/c
Sequence 49730, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49730
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-49730

Query Match 45.5%; Score 23.2; DB 4; Length 601;
Best Local Similarity 70.5%; Pred. No. 44;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 4 GGGGCACTCCGCTGCTCTGACTGCGCCCGCAGTGA 47
DB 453 GAGGCACTGCGCTCCCTTGAATTGAGCTCAGGCTTCAG 410

RESULT 15
US-09-949-016-49731/c

```
/ Sequence 49731, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 49731
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-49731

Query Match      45.5%; Score 23.2; DB 4; Length 601;
Best Local Similarity 70.5%; Pred. No. 44;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      4 GGGCGACGGGGGACTCCCGTGTCTCTGAGACTCTGGCCCCCAG 47
         |||||
Db      452 GAGCCAACCGCGCCCTCCCTGTGATCTTGAGCTCCAGCCTCCAG 409
```

Search completed: September 1, 2005, 07:07:03
Job time : 22.2736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 31, 2005, 12:52:03 ; Search time 286.304 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916B-3

Perfect score: 270

Sequence: 1 catggcgacgacggcgaccc.....ttccgctactagctagcgcg 270

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980a:*
2: geneseqn1980b:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	100.0	270	6	ABX15373 Rat insul
2	270	100.0	423	6	ABX15376 Rat insul
3	219.4	81.3	321	6	ABX15374 Rat insul
4	219.4	81.3	372	6	ABX15375 Rat insul
5	219	81.1	219	6	ABX15372 Rat insul
6	76.4	28.3	1500	6	ABK63701 Rat seque
7	76.4	28.3	1500	10	ADBS8201
8	76.4	28.3	1500	10	ADBS2710
9	76.4	28.3	1500	10	ABR41911
10	76.4	28.3	1500	12	ADP72689
11	76.4	28.3	5001	4	AAH22429
12	60.8	22.5	6128	6	ABN95896
13	60.8	22.5	6128	10	ABV75371
14	60.8	22.5	6134	4	AAH57489
15	51	18.9	51	6	ABX15371
16	51	18.9	194	2	AAH33002
17	51	18.9	321	6	ABX15374 Rat insul
18	51	18.9	372	6	ABX15375 Rat insul
19	51	18.9	13011	2	AAH96631
20	51	18.9	13011	10	ABT42448

21	51	18.9	13011	12	ADP72914	Adp72914 Renal tox
22	48	17.8	48	6	ABX15379	Abx15379 Rat liver
23	48	17.8	51	6	ABX15380	Abx15380 Rat liver
24	37.8	14.0	7061	6	ABL32997	Abi32997 Human imm
25	37.8	14.0	7061	6	ABL70248	Abi70248 Chemical1
26	37.8	14.0	7061	6	AAH61189	AAH61189 Human gen
27	35	13.0	3895	6	ABN87725	Abn87725 Human pro
28	33.4	12.4	1806	6	ABO75795	Abq75795 Human MDD
29	33.4	12.4	2667	11	ADM03554	Adm03554 Human GDN
30	33.2	12.3	650	8	ABH52705	Abh52705 Apep911
31	32.4	12.0	1048	5	AAH69850	AAH69850 DNA encod
32	32.2	11.9	4447	2	AAH14078	AAH14078 pig dthyd
33	32	11.9	76180	13	ABD33385	Abd33385 Human can
34	31.6	11.7	548	13	ADQ051481	Adq51481 Novel can
35	31.6	11.7	927	6	ABH57730	Abh57730 Human bbg
36	31.6	11.7	1125	6	AAH94440	AAH94440 Human hyd
37	31.6	11.7	1129	6	ADJ33646	Adj33646 Human trf
38	31.6	11.7	1288	10	ADC87210	Adc87210 Human GPC
39	31.6	11.7	1344	6	ABH57731	Abh57731 Human bbg
40	31.6	11.7	1348	8	ABX71176	Abx71176 Novel hum
41	31.6	11.7	1513	6	ADH16493	Adh16493 Human NOV
42	31.6	11.7	1550	4	AAH94450	AAH94450 Human hyd
43	31.6	11.7	1555	4	AAH90027	AAH90027 Human dig
44	31.6	11.7	1555	5	AAH39684	AAH39684 Genomic B
45	31.6	11.7	1555	9	ADB32644	Adb32644 Human nov

ALIGNMENTS

RESULT 1	ABX15373	standard; DNA, 270 BP.
ID	ABX15373	
XX	ABX15373;	
AC	XX	
XX	XX	
DT	17-APR-2003	(first entry)
XX	XX	
DE	Rat insulin regulator construct DNA #1.	
KW	Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGF1P-1;	
KW	insulin-sensitive element; ISE; basal promoter; hyperglycemia; insulin;	
KW	insulin-like growth factor binding protein-1; hypoglycemia; glucose;	
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;	
KW	glucagon; euglycemia; diabetes; fasting; ketogenesis; ketoacidosis;	
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;	
KW	intracellular hormone receptor; insulin regulator construct; anabolic.	
XX	XX	
OS	Rattus norvegicus.	
OS	Synthetic.	
XX	XX	
PN	US2002107198-A1.	
XX	XX	
PD	08-AUG-2002.	
XX	XX	
PF	10-OCT-2001; 2001US-00972916.	
XX	XX	
PR	11-OCT-2000; 2000US-0239113P.	
XX	XX	
PA	(THUL/) THUL P M.	
XX	XX	
PI	Thule PM;	
XX	XX	
DR	WPI; 2002-674190/72.	
XX	XX	
PT	New insulin regulator cassette, useful e.g. for treating diabetes,	
PT	provides specific, glucose-inducible transgenic expression of insulin in	
XX	liver cells.	
PS	Claim 9; Page 13-14; 37pp; English.	
XX	XX	
CC	The invention relates to an insulin regulator construct comprising a	
CC	glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene	

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 270; DB 6; Length 270;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 DB 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 QY 61 AAACAACTTATTTTGAACACGGGGATCCTAGACACCTGCGCTGACATCATTAACCGCT 120
 DB 61 AAACAACTTATTTTGAACACGGGGATCCTAGACACCTGCGCTGACATCATTAACCGCT 120
 QY 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGTCCACATGGCC 180
 DB 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGTCCACATGGCC 180
 QY 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACAGCGCATCTGCCAGAGAGCTG 240
 DB 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACAGCGCATCTGCCAGAGAGCTG 240
 QY 241 TGACCAACCACTTCCGCTACTAGTACGCCG 270
 DB 241 TGACCAACCACTTCCGCTACTAGTACGCCG 270

RESULT 2

ABX15376
 ID ABX15376 standard; DNA; 423 BP.

AC ABX15376;
 XX

DT 17-APR-2003 (first entry)
 XX

DE Rat insulin regulator construct DNA #4.
 XX

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucoacetic acid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Ratus norvegicus.
 OS Synthetic.
 XX

PN US2002107198-A1.
 XX

PD 08-AUG-2002.
 XX

PF 10-OCT-2001; 2001US-00972916.
 XX

PR 11-OCT-2000; 2000US-0239113P.
 XX

XX (THUL/) THULE P M.
 PA
 XX
 PI Thule PM;
 DR WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.

XX Claim 9; Page 14; 37pp; English.

XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 270; DB 6; Length 423;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 DB 154 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 213
 QY 61 AAACAACTTATTTTGAACACGGGGATCCTAGACACCTGCGCTGACATCATTAACCGCT 120
 DB 214 AAACAACTTATTTTGAACACGGGGATCCTAGACACCTGCGCTGACATCATTAACCGCT 273
 QY 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGTCCACATGGCC 180
 DB 274 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGTCCACATGGCC 333
 QY 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACAGCGCATCTGCCAGAGAGCTG 240
 DB 334 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACAGCGCATCTGCCAGAGAGCTG 393
 QY 241 TGACCAACCACTTCCGCTACTAGTACGCCG 270
 DB 394 TGACCAACCACTTCCGCTACTAGTACGCCG 423

RESULT 3

ABX15374
 ID ABX15374 standard; DNA; 321 BP.

AC ABX15374;
 XX

DT 17-APR-2003 (first entry)
 XX

DE Rat insulin regulator construct DNA #2.
 XX

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;

KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
OS Rattus norvegicus.
OS Synthetic.
PN US2002107198-A1.
XX
XX
XX 08-AUG-2002.
PD
XX
XX 10-OCT-2001; 2001US-00972916.
PF
XX
XX 11-OCT-2000; 2000US-0239113P.
PR
XX
XX (THUL/) THULE P M.
PA
XX
XX Thule PM;
PI
XX
XX WPI; 2002-674190/72.
DR
XX
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
PS
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX
SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
Query Match 81.3%; Score 219.4; DB 6; Length 321;
Best Local Similarity 99.5%; Pred. No. 8.7e-61;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 50 TATCAACAAGCAAAACAATTTGAAACGGGGATCTTGAACAGCGTCCCTGACAT 109
DB 101 TGTCAACAACAAACAACTTATTGAAACGGGGATCTTGAACAGCGTCCCTGACAT 160
QY 110 CATTAAACCGGTGCTGCGAGACGAGCCCTTCATTAAGGCGCTGGATATGCGCAGCAGCATG 169
DB 161 CATTAAACCGGTGCTGCGAGACGAGCCCTTCATTAAGGCGCTGGATATGCGCAGCAGCATG 220
QY 170 GTTCACTGCGCGCGAGACCAAAACCCAGCAGCATTTGAACACTGCAACGGCCATCTGC 229
DB 221 GTTCACTGCGCGCGAGACCAAAACCCAGCAGCATTTGAACACTGCAACGGCCATCTGC 280
QY 230 CCAGAGAGCTGTGACGACCACTTCGCTACTAGCTGCGCGC 270
DB 281 CCAGAGAGCTGTGACGACCACTTCGCTACTAGCTGCGCGC 321

RESULT 4
ABX15375

ID ABX15375 strand; DNA; 372 BP.
XX
XX AC ABX15375;
XX
XX 17-APR-2003 (first entry)
XX
XX DE Rat insulin regulator construct DNA #3.
XX
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX
XX Rattus norvegicus.
OS
OS Synthetic.
PN US2002107198-A1.
XX
XX
XX 08-AUG-2002.
PD
XX
XX 10-OCT-2001; 2001US-00972916.
PF
XX
XX 11-OCT-2000; 2000US-0239113P.
PR
XX
XX (THUL/) THULE P M.
PA
XX
XX Thule PM;
PI
XX
XX WPI; 2002-674190/72.
DR
XX
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
PS
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX
SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
Query Match 81.3%; Score 219.4; DB 6; Length 372;
Best Local Similarity 99.5%; Pred. No. 9.2e-61;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 50 TATCAACAAGCAAAACAATTTGAAACGGGGATCTTGAACAGCGTCCCTGACAT 109
DB 152 TGTCAACAAGCAAAACAATTTGAAACGGGGATCTTGAACAGCGTCCCTGACAT 211
QY 110 CATTAAACCGGTGCTGCGAGACGAGCCCTTCATTAAGGCGCTGGATATGCGCAGCAGCATG 169
DB 212 CATTAAACCGGTGCTGCGAGACGAGCCCTTCATTAAGGCGCTGGATATGCGCAGCAGCATG 271

```
QY 170 GTCCACTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGCACAGGCCATCTGC 229
DB 272 GTCCACTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGCACAGGCCATCTGC 331
QY 230 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270
DB 332 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 372

RESULT 5
ABK15372
ID ABK15372 standard; DNA; 219 BP.
AC ABK15372;
XX
XX 17-APR-2003 (first entry)
XX
XX Rat insulin-sensitive element (ISE) DNA.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 8; Page 13; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin-sensitive element of the invention
XX
XX Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;
XX
XX Query Match 81.1%; Score 219; DB 6; Length 219;
XX Best Local Similarity 100.0%; Pred. No. 1e-60;
```

```
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 52 TCACAGAGAAAACAACCTTATTGGAACAGGGATCTTACAGCTCCCTGACAAATCA 111
DB 1 TCACAGAGAAAACAACCTTATTGGAACAGGGATCTTACAGCTCCCTGACAAATCA 60
QY 112 TTAACCCGTGCTGCGAGACCAAGCCCTTCATTAAGCCCTGGGTATGAGCCAGCATGTGT 171
DB 61 TTAACCCGTGCTGCGAGACCAAGCCCTTCATTAAGCCCTGGGTATGAGCCAGCATGTGT 120
QY 172 CCACCTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGCACAGGCCATCTGC 231
DB 121 CCACCTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGCACAGGCCATCTGC 180
QY 232 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219

RESULT 6
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
AC ABK63701;
XX
XX 18-JUN-2002 (first entry)
XX
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
XX
XX WO200210453-A2.
XX
XX 07-FEB-2002.
XX
XX 30-JUL-2001; 2001WO-US023872.
XX
XX 31-JUL-2000; 2000US-0222040P.
XX 02-NOV-2000; 2000US-0244880P.
XX 11-MAY-2001; 2001US-0290029P.
XX 15-MAY-2001; 2001US-0290645P.
XX 22-MAY-2001; 2001US-0292336P.
XX 06-JUN-2001; 2001US-0295798P.
XX 13-JUN-2001; 2001US-0297457P.
XX 19-JUN-2001; 2001US-0298884P.
XX 09-JUL-2001; 2001US-0303459P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
```

CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilize a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridizes to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity is
 CC characterized by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent
 CC
 XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 6; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 247
 DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 68 CACTTCGGCTACTAGCTA 85

RESULT 7
 ADB58201
 ID ADB58201 standard; DNA; 1500 BP.

AC ADB58201;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3227.
 XX
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 XX drug screening; toxicity assay; ds.
 XX
 OS Unidentified.
 XX
 PN WO2003064624-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 31-JAN-2003; 2003WO-US003194.
 XX
 PR 31-JAN-2002; 2002US-00060087.
 PR 15-MAR-2002; 2002US-0364045P.
 PR 15-MAR-2002; 2002US-0364055P.
 PR 30-DEC-2002; 2002US-0436643P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 XX
 DR MPI; 2003-689530/65.
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 PT markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 PT the compound.
 XX

PS Claim 1; SEQ ID NO 3227; 1156pp; English.

XX The present invention relates to a method for predicting a toxic effect
 CC of a compound. The method comprises preparing a gene expression profile
 CC of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 CC of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 247
 DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 68 CACTTCGGCTACTAGCTA 85

RESULT 8
 ADB52710
 ID ADB52710 standard; DNA; 1500 BP.

AC ADB52710;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3852.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 XX toxicity marker; toxicity progression; drug screening;
 XX primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003065993-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003WO-US003482.
 XX
 PR 04-FEB-2002; 2002US-0353171P.
 PR 13-MAR-2002; 2002US-036534P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.
 PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373601P.
 PR 19-APR-2002; 2002US-0373602P.
 PR 22-APR-2002; 2002US-0374139P.
 PR 08-MAY-2002; 2002US-0378370P.
 PR 09-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-JUL-2002; 2002US-0394230P.
 PR 09-JUL-2002; 2002US-0394253P.

```

PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442300P.
XX
PA (GENE-) GENE LOGIC INC.
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3252; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiologic state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
Query Match 28.3%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.4e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 247
DB 8 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
XX
QY 248 CACTTCGGCTACTAGCTA 265
DB 68 CACTTCGGCTACTAGCTA 85
XX
RESULT 9
ABTA1911
ID ABTA1911 standard; DNA; 1500 BP.
XX
AC ABTA1911;
XX
DT 26-JUN-2003 (first entry)
XX
DE Toxicity modelling related rat gene SEQ ID No 1613.
XX
KM Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KM database; drug screening; toxicity assay; rat; ds.
XX
OS Rattus norvegicus.
XX
PN WO200295000-A2.
XX
PD 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
PR 22-MAY-2001; 2001US-0292335P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
XX

```

```

PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357843P.
PR 15-MAR-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX WPI; 2003-148464/14.
XX
PT Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
PS Example 4; Page; 446pp; English.
XX
CC The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
Query Match 28.3%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.4e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 247
DB 8 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
XX
QY 248 CACTTCGGCTACTAGCTA 265
DB 68 CACTTCGGCTACTAGCTA 85
XX
RESULT 10
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
AC ADP72689;
XX
DT 26-AUG-2004 (first entry)
XX
DE Renal toxin progression gene marker #1278.
XX

```


KW db; toxic effect; gene expression profile; kidney tissue;
KW differential gene expression; toxicity progression; toxicity marker;
KW drug screening; toxicity assay; kidney pathology; nephritis;
KW kidney necrosis; glomerular injury; tubular injury;
KW focal segmental glomerulosclerosis.
OS Rattus norvegicus.
XX
XX WO2004048598-A2.
XX
XX 10-JUN-2004.
XX
XX 24-NOV-2003; 2003WO-US037556.
XX
XX 22-NOV-2002; 2002US-00301856.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
PI Elashoff M;
XX
XX WPI; 2004-460771/43.
XX
XX
XX Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
PS
XX Claim 11; SEQ ID NO 1278; 266pp; English.
XX
XX The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.4e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGACACTGACACGGCCATCTGCCAGAGAGCTGTACAC 247
DB 8 CACAAACCCAGGAGCATTTGACACTGACACGGCCATCTGCCAGAGAGCTGTACAC 67
QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 11
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX

AC AAH22429;
XX
XX 22-AUG-2001 (first entry)
XX
XX
XX Rat insulin-like growth factor binding protein nucleotide sequence.
DE
XX
XX Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW non-steroidal antiinflammatory drug; ds.
KW
XX
XX Rattus norvegicus.
OS
XX
XX WO200138579-A2.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US032049.
XX
XX 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
XX (CURA-) CURAGEN CORP.
XX
XX
XX Gould-Rothberg BE, Dipippo VA, Rameeh TM, Gerwein RW;
PI
XX
XX WPI; 2001-355948/37.
XX
XX
XX Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
PS
XX Disclosure; Page 22-24; 76pp; English.
XX
XX The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in a reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention.
XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 5.3e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGACACTGACACGGCCATCTGCCAGAGAGCTGTACAC 247
DB 1 CACAAACCCAGGAGCATTTGACACTGACACGGCCATCTGCCAGAGAGCTGTACAC 60
QY 248 CACTTCCGCTACTAGCTA 265
DB 61 CACTTCCGCTACTAGCTA 78

RESULT 12
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
XX ABN95896;
XX
XX 13-AUG-2002 (first entry)
XX

Best Local Similarity 66.5%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTACGACGCTGCCCTGACATCAT 112
DB 473 CACTACCAAAACAACTTATTTTGAACACGCTCTTACGACGCTGCCCTGACATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAGGCGCTGGGTATGGCCAGC 162
DB 533 TAACCTCTCGTGCAAGTGGCGGCGGCTGTGCTTATTAAGGTGGCGCTGTGTCCAGC 592
QY 163 CAGCATGTCTCACTGCGCCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 221
DB 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCAGCATCTCCCGCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCACTTC 253
DB 645 CCACCTCCGAGAGACACTGGCCACCGCTCC 676

RESULT 14

AAH57489
ID AAH57489 standard; cDNA; 6134 BP.

AAH57489;
10-SEP-2001 (first entry)

Human liver cell specific cDNA sequence SEQ ID NO:329.

Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
metabolic disease; developmental disease; cytosolic; immunomodulatory;
neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX Homo sapiens.
XX OS
XX PN MO200132927-A2.
XX PD 10-MAY-2001.
XX PF 02-NOV-2000; 2000MO-US030396.
XX PR 04-NOV-1999; 99US-0163508P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Sornasse T, Seilhamer J, Watson GA;
XX MPI; 2001-291057/30.
XX DR
XX PT New cell and tissue specific polynucleotides useful for diagnosis,
XX PT prognosis or monitoring of treatments for disorders where the gene is
XX PT associated with a cancer, immunopathology or neuropathology.
XX PS
XX PS Claim 1; Page 246-248; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
sequences (I). (I) can have cytosolic, immunomodulatory and
neuroprotective activities, and can be used in gene therapy. (I) and
proteins (II) encoded by them are used in high throughput screening
assays to select DNA molecules, RNA molecules, peptide nucleic acids,
mimetics, peptides, proteins, agonists, antagonists, antibodies or their
fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
agents. Expression of (I) in a sample indicates the differentiation of
embryonic stem cells into a tissue selected from brain, heart, kidney,
liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
to produce an expression profile that defines a metabolic or
developmental process, treatment, condition, disease or disorder. The
gene profile can be used for diagnosis, prognosis or monitoring of
treatments and for investigating a predisposition to a disorder where the
gene is associated with a cancer, immunopathology or neuropathology

Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;

Query Match 22.5%; Score 60.8; DB 4; Length 6134;
Best Local Similarity 66.5%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTACGACGCTGCCCTGACATCAT 112
DB 473 CACTACCAAAACAACTTATTTTGAACACGCTCTTACGACGCTGCCCTGACATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAGGCGCTGGGTATGGCCAGC 162
DB 533 TAACCTCTCGTGCAAGTGGCGGCGGCTGTGCTTATTAAGGTGGCGCTGTGTCCAGC 592
QY 163 CAGCATGTCTCACTGCGCCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 221
DB 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCAGCATCTCCCGCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCACTTC 253
DB 645 CCACCTCCGAGAGACACTGGCCACCGCTCC 676

RESULT 15

ABX15371
ID ABX15371 standard; DNA; 51 BP.

ABX15371;

17-APR-2003 (first entry)

Rat glucose response element (GIRE) DNA.

Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
insulin-sensitive element; ISR; basal promoter; hyperglycaemia; insulin;
insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
intracellular hormone receptor; insulin regulator construct; anabolic.
XX Rattus norvegicus.
XX OS
XX PN US2002107198-A1.
XX PD 08-AUG-2002.
XX PF 10-OCT-2001; 2001US-00972916.
XX PR 11-OCT-2000; 2000US-0239113P.
XX PA (THUL/) THULE P M.
XX PI Thule PM;
XX DR MPI; 2002-674190/72.

New insulin regulator cassette, useful e.g. for treating diabetes,
provides specific, glucose-inducible transgene expression of insulin in
liver cells.

Claim 7; Page 13; 37pp; English.

The invention relates to an insulin regulator construct comprising a
glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
promoter and an insulin-sensitive element (ISR) of an insulin-like growth
factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
to treat or prevent diabetic complications, to regulate insulin
production, to modulate hyperglycaemia (without severe hypoglycaemia), to
increase fat catabolism and to retard protein catabolism. The construct
provides stimulation of insulin expression by glucose (but not other
carbohydrates) and glucocorticoids and inhibition by glucagon so that a
combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat glucose response element of the invention
 XX

SQ Sequence 51 bp; 6 A; 18 C; 17 G; 10 T; 0 U; 0 Other;

Query Match

18.9%; Score 51; DB 6; Length 51;

Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGGACGCGGGGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGTA 51
 |||||||
 DB 1 CATGGCGGACGCGGGGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGTA 51

Search completed: September 1, 2005, 00:34:43
 Job time : 288.304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 31, 2005, 12:52:03 ; Search time 340.384 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321

Sequence: 1 taacactgggggccaagatcc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002as:*
8: geneseqn2003as:*
9: geneseqn2003as:*
10: geneseqn2003as:*
11: geneseqn2003as:*
12: geneseqn2004as:*
13: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	321	100.0	321	6	ABX15374 Rat insul
2	321	100.0	372	6	ABX15375 Rat insul
3	219.4	68.3	270	6	ABX15373 Rat insul
4	219.4	68.3	423	6	ABX15376 Rat insul
5	219	68.2	219	6	ABX15372 Rat insul
6	103	32.1	423	6	ABX15376 Rat insul
7	76.4	23.8	1500	6	ABK63701 Rat segue
8	76.4	23.8	1500	10	ADBS8201 Toxicity-
9	76.4	23.8	1500	10	ADBS52710 Primary r
10	76.4	23.8	1500	10	ABR41911 Toxicity
11	76.4	23.8	1500	12	ADP72689 Renal tox
12	76.4	23.8	5001	4	AAH22429 Rat insul
13	63.8	19.9	6128	6	ABN95896 Gene #239
14	63.8	19.9	6128	10	ABV75371 Human IGF
15	63.8	19.9	6134	4	AAH57489 Human liv
16	55.2	17.2	1394	2	AAH57489 Human liv
17	55.2	17.2	13011	2	AAH57489 Human liv
18	55.2	17.2	13011	10	ABR42448 Toxicity
19	55.2	17.2	13011	12	ADP72914 Renal tox
20	51	15.9	51	6	ABX15380 Rat liver

ALIGNMENTS

RESULT 1	ABX15374	standard; DNA; 321 BP.
ID	ABX15374	standard; DNA; 321 BP.
AC	ABX15374;	
XX		
DT	17-APR-2003	(first entry)
XX		
DE	Rat insulin regulator construct DNA #2.	
XX		
KW	Glucose response element; GIRE; liver pyruvate kinase; Lpk; ds; IGF-1;	Abx15371 Rat gluco
KW	insulin-sensitive element; ISE; basal promoter; hyperglycemia; insulin;	Abx15373 Rat insul
KW	insulin-like growth factor binding protein-1; hypoglycemia; glucose;	Abx15379 Rat liver
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;	Ab132997 Human imm
KW	glucagon; euglycemia; diabetes; fasting; ketogenesis; ketoacidosis;	Ab170248 Human imm
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;	Ab661189 Human gen
KW	intracellular hormone receptor; insulin regulator construct; anabolic.	AdA52441 Human cod
XX		
OS	Rattus norvegicus.	AbB70481 Human bon
OS	Synthetic.	Ab190266 Human pol
OS		Ab190266 Human pol
PN	US2002107198-A1.	Ab190266 Human pol
XX		
PD	08-AUG-2002.	Ab190266 Human pol
XX		
PF	10-OCT-2001; 2001US-00972916.	Ab190266 Human pol
XX		
PR	11-OCT-2000; 2000US-0239113P.	Ab190266 Human pol
XX		
PA	(THUL/) THUL F M.	Ab190266 Human pol
XX		
PI	Thule PM;	Ab190266 Human pol
XX		
DR	WPI; 2002-674190/72.	Ab190266 Human pol
XX		
PT	New insulin regulator cassette, useful e.g. for treating diabetes,	Ab190266 Human pol
XX	provides specific, glucose-inducible transgenic expression of insulin in	Ab190266 Human pol
PT	liver cells.	Ab190266 Human pol
XX		
PS	Claim 9; Page 14; 37pp; English.	Ab190266 Human pol
XX		
CC	The invention relates to an insulin regulator construct comprising a	Ab190266 Human pol
XX	glucose response element (GIRE) of a liver pyruvate kinase (Lpk) gene	Ab190266 Human pol

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 100.0%; Score 321; DB 6; Length 321;

Best Local Similarity 100.0%; Pred. No. 4, 4e-89;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCACTGGGGGCGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 Db 1 TCACTGGGGGCGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 QY 61 GGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 120
 Db 61 GGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 120
 QY 121 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 180
 Db 121 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 180
 QY 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCCCGCAGACA 240
 Db 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCCCGCAGACA 240
 QY 241 CAAACCCAGCGAGCATTTGAACATCTGCACACGGCCATCTGCCAGAGAGTGTGACACCA 300
 Db 241 CAAACCCAGCGAGCATTTGAACATCTGCACACGGCCATCTGCCAGAGAGTGTGACACCA 300
 QY 301 CTTCCGCTACTAGTACCGGC 321
 Db 301 CTTCCGCTACTAGTACCGGC 321

RESULT 2

ABX15375
 ID ABX15375 standard; DNA; 372 BP.

AC ABX15375;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #3.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
 OS Synthetic.

XX US2002107198-A1.

XX 08-AUG-2002.
 PD

XX 10-OCT-2001; 2001US-00972916.
 PF 11-OCT-2000; 2000US-0239113P.
 PR (THUL/) THULE P M.
 XX
 PA Thule PM;
 PI WPI; 2002-674190/72.
 DR

XX New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.

PS Claim 9; Page 14; 37pp; English.

XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 321; DB 6; Length 372;

Best Local Similarity 100.0%; Pred. No. 4, 7e-89;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCACTGGGGGCGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 Db 52 TCACTGGGGGCGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 111
 QY 61 GGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 120
 Db 112 GGCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 171
 QY 121 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 180
 Db 121 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 180
 QY 172 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 231
 Db 172 TATTTGAACAGGGGATCTCTAGCAGCTGCCCTGCATTCATTACCGGTGTCGGAG 231
 QY 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCCCGCAGACA 240
 Db 232 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCCCGCAGACA 291
 QY 241 CAAACCCAGCGAGCATTTGAACATCTGCACACGGCCATCTGCCAGAGAGTGTGACACCA 300
 Db 292 CAAACCCAGCGAGCATTTGAACATCTGCACACGGCCATCTGCCAGAGAGTGTGACACCA 351
 QY 301 CTTCCGCTACTAGTACCGGC 321
 Db 352 CTTCCGCTACTAGTACCGGC 372

RESULT 3

ABX15373

ID ABX15373 standard; DNA; 270 BP.

XX ABX15373;
 AC

```

XX 17-APR-2003 (first entry)
DT Rat insulin regulator construct DNA #1.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THUL P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 9; Page 13-14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis.
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
XX
XX Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;
SQ
XX
XX Query Match 68.3%; Score 219.4; DB 6; Length 270;
XX Best Local Similarity 99.5%; Pred. No. 1.1e-57;
XX Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 101 TGTGACAGCAAAACAACTTATTTGAAACGCGGATCTTAGACAGCTGCCCTGACAAAT 160
XX DB 50 TATGACAAACAAACAACTTATTTGAAACGCGGATCTTAGACAGCTGCCCTGACAAAT 109
XX
XX 161 CATTACCCGCTGCTGCGGAGCCAGCCCTTCATTAAGCCCTCGGTATAGGCAACGACAG 220
XX DB 110 CATTAAACCGTGTGCTGCGGAGCCAGCCCTTCATTAAGCCCTCGGTATAGGCAACGACAG 169
XX
XX 221 GTTCACTGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGCACAGGCAATCTGC 280
XX DB 170 GTTCACTGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGCACAGGCAATCTGC 229

```

```

XX 281 CCAGAGAGCTGTGACACCACTTCGCGTACTAGTACGCGC 321
XX DB 230 CCAGAGAGCTGTGACACCACTTCGCGTACTAGTACGCGC 270
XX
XX RESULT 4
XX ID ABX15376 standard; DNA; 423 BP.
XX
XX ABX15376;
XX
XX 17-APR-2003 (first entry)
DT Rat insulin regulator construct DNA #4.
XX
XX Rat insulin regulator construct DNA #4.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THUL P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
XX
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
SQ
XX
XX Query Match 68.3%; Score 219.4; DB 6; Length 423;
XX Best Local Similarity 99.5%; Pred. No. 1.3e-57;
XX Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetic during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis.
 CC Thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
 Query Match 32.1%; Score 103; DB 6; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.2e-21;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACACTGGGGGCCAGAGTCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 60
 DB 204 TACACTGGGGGCCAGAGTCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 145
 QY 61 GGGCAGAGTCCAGAGCAACGAGGAGTCCCGTCCGCCATGT 103
 DB 144 GGGCAGAGTCCAGAGCAACGAGGAGTCCCGTCCGCCATGT 102
 RESULT 7
 ABK63701
 ID ABK63701 standard; cDNA; 1500 BP.
 AC ABK63701;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
 XX
 KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KM differential expression; centrilobular necrosis; steatosis.
 OS Rattus norvegicus.
 XX
 XX WO200210453-A2.
 XX
 XX 07-FEB-2002.
 XX
 PD 30-JUL-2001; 2001WO-US023872.
 PF
 XX 31-JUL-2000; 2000US-0222040P.
 XX PR 02-NOV-2000; 2000US-0244880P.
 XX PR 11-MAY-2001; 2001US-0290029P.
 XX PR 15-MAY-2001; 2001US-0290645P.
 XX PR 22-MAY-2001; 2001US-0292336P.
 XX PR 06-JUN-2001; 2001US-0295798P.
 XX PR 13-JUN-2001; 2001US-0297457P.
 XX PR 19-JUN-2001; 2001US-0298884P.
 XX PR 09-JUL-2001; 2001US-0303459P.
 XX
 XX (GENE-) GENE LOGIC INC.
 XX
 XX Mendrick D, Porter WM, Johnson KR, Caesle AL, Elashoff MR;
 XX
 DR MPI; 2002-241625/29.
 XX
 XX Predicting toxic effects of compounds or the progression of these toxic
 PT effects by determining the changes in gene expression in tissues or cells
 PT exposed to the toxin and comparing these to gene expression in unexposed
 PT tissues or cells.
 XX
 PS Claim 1; SEQ ID NO 1608; 239pp; English.

XX The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic effect
 CC of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression. The
 CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilize a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridizes to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity is
 CC characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent
 CC
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 6; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTGAACACTGCACAGGCGCATCTGCCAGAGCTGTGACCAC 298
 DB 8 CACAAACCCAGGAGCATTGAACACTGCACAGGCGCATCTGCCAGAGCTGTGACCAC 67
 QY 299 CACTTCCGCTACTACTA 316
 DB 68 CACTTCCGCTACTACTA 85
 RESULT 8
 ADB58201
 ID ADB58201 standard; DNA; 1500 BP.
 AC ADB58201;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3227.
 XX
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 KM drug screening; toxicity assay; de.
 XX
 OS Unidentified.
 XX
 XX WO2003064624-A2.
 XX
 XX 07-AUG-2003.
 XX
 PD 31-JAN-2003; 2003WO-US003194.
 PF
 XX 31-JAN-2002; 2002US-00060087.
 XX PR 15-MAR-2002; 2002US-0364045P.
 XX PR 15-MAR-2002; 2002US-0364055P.
 XX PR 30-DEC-2002; 2002US-0436643P.
 XX

PA (GENE-) GENE LOGIC INC.
 XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 PI WPI, 2003-689530/65.
 XX
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 the compound.
 XX
 XX
 PS Claim 1; SEQ ID NO 3227, 1156bp; English.
 XX
 XX The present invention relates to a method for predicting a toxic effect
 of a compound. The method comprises preparing a gene expression profile
 of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
 DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 299 CACTTCGCTACTATCTA 316
 DB 68 CACTTCGCTACTATCTA 85
 DB
 RESULT 9
 ID ADB52710 standard; DNA; 1500 BP.
 XX
 AC ADB52710;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
 XX
 XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003065993-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003WO-US003482.
 XX
 PR 04-FEB-2002; 2002US-0353171P.
 PR 13-MAR-2002; 2002US-036534P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373601P.
 PR 19-APR-2002; 2002US-0373602P.
 PR 22-APR-2002; 2002US-0374139P.
 PR 08-MAY-2002; 2002US-0378370P.
 PR 09-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-JUL-2002; 2002US-0394230P.
 PR 09-JUL-2002; 2002US-0394253P.
 PR 04-SEP-2002; 2002US-0407688P.
 PR 28-JAN-2003; 2003US-0442900P.
 XX
 XX (GENE-) GENE LOGIC INC.
 XX
 XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
 PI Elashoff M;
 XX WPI, 2003-731472/69.
 XX
 DR
 XX
 PT Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
 PT mean values.
 XX
 PS Claim 44; SEQ ID NO 3252; 874bp; English.
 XX
 XX The present invention describes a method for determining whether a
 CC compound induces a toxic effect on a tissue or cell. The method comprises
 CC preparing a gene expression profile of a tissue or cell sample exposed to
 CC the compound, and comparing the gene expression profile to a database
 CC comprising data or information on the Tox mean and non-Tox mean value.
 CC The method is useful for predicting or identifying at least one toxic
 CC effect, particularly hepatotoxicity, of a test or unknown compound. The
 CC genes listed in the specification are useful as diagnostic or toxicity
 CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
 DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 299 CACTTCGCTACTATCTA 316
 DB 68 CACTTCGCTACTATCTA 85
 DB
 RESULT 10
 ID ABT41911 standard; DNA; 1500 BP.
 XX
 AC ABT41911;
 XX
 DT 26-JUN-2003 (first entry)
 XX
 DE Toxicity modelling related rat gene SEQ ID No 1613.
 XX
 XX Toxic effect; gene expression profile; renal toxicity; toxicity marker;
 KW database; drug screening; toxicity assay; rat; ds.
 XX
 OS Rattus norvegicus.
 XX

PN WO200295000-A2.
XX
XX
PD 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
PR 12-MAY-2001; 2001US-0292335P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2002US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M,
XX
XX WPI; 2003-148464/14.
DR
XX
XX
PT Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX
PS Example 4; Page; 446pp; English.
XX
XX
CC The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the renal
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 23.8%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.2e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
XX
AC ADP72689;
XX
XX 26-AUG-2004 (first entry)
XX
XX Renal toxin progression gene marker #1278.
DE
XX
XX ds; toxic effect; gene expression profile; kidney tissue;
XX differential gene expression; toxicity progression; toxicity marker;
XX drug screening; toxicity assay; kidney pathology; nephritis;
XX kidney necrosis; glomerular injury; tubular injury;
XX focal segmental glomerulosclerosis.
XX
XX Rattus norvegicus.
OS
XX
XX WO2004048598-A2.
XX
XX 10-JUN-2004.
XX
XX 24-NOV-2003; 2003WO-US037556.
PF
XX
XX 22-NOV-2002; 2002US-00301856.
PR
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
XX
XX Elashoff M;
XX
XX WPI; 2004-460771/43.
DR
XX
XX
PT Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
XX
PS Claim 11; SEQ ID NO 1278; 266pp; English.
XX
XX
XX
CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC monitoring or modelling the toxic response of a test compound, for
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 23.8%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.2e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGGAGATTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACAC 298
|||||
8 CACAAACCCAGGAGATTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACAC 67
|||||
QY 239 CACTTCGGCTACTAGCTA 316
|||||
DB 68 CACTTCGGCTACTATCTA 85
|||||

Db 8 CACAAACCCAGGACATTGAACTGACACAGCCATCTGCCAGAGCTGTGACAC 67
OY 239 CACTTCGGCTACTAGCTA 316
|||
Db 68 CACTTCGGCTACTAGCTA 85
|||
RESULT 12
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
AC AAH22429;
XX
DT 22-AUG-2001 (first entry)
XX
DE Rat insulin-like growth factor binding protein nucleotide sequence.
KM Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW non-steroidal antiinflammatory drug; ds.
XX
OS Rattus norvegicus.
XX
PN WO200138579-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032049.
XX
PR 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gould-Rothberg BE, Dipippo VA, Rameeh TM, Gerwein RW;
XX
XX WPI; 2001-355948/37.
XX
PT Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
PS Disclosure; Page 22-24; 76pp; English.
XX
CC The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in an reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;
Query Match 23.8%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 4.7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 239 CACAAACCCAGGACATTGAACTGACACAGCCATCTGCCAGAGCTGTGACAC 298
|||
Db 1 CACAAACCCAGGACATTGAACTGACACAGCCATCTGCCAGAGCTGTGACAC 60
|||
OY 239 CACTTCGGCTACTAGCTA 316
|||

Db 61 CACTTCGGCTACTAGCTA 78
RESULT 13
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
XX ABN95896;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2394 used to diagnose liver cancer.
XX
KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumour; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-da-Silva S, Vockley JG;
XX
XX WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 2394; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cyrostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
Query Match 19.9%; Score 63.8; DB 6; Length 6128;
Best Local Similarity 63.2%; Pred. No. 3.9e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
OY 69 TCCAGGAACCCAGGAGTGCCTCGGCCCATGTCAAGCAAAACCTTATTTTGA 128
|||
Db 438 TCCATCCACACAGCGGTTTGCTAGAGGCTTGGGTCACCTAGCAAAACCTTATTTTGA 497
|||
OY 129 ACAGCGGAGATTGACAGCGCTGCCCTGACATCTTAACCC-----GTCTGCGG 178
|||
Db 498 ACACTCAGCTTCCTAGCGCTGCGCGCTCTCCATCTTAACCTCTGTGCAAGTGCGCGG 557
|||
OY 179 AGCCAGCCCTTCATAAAGCCCTGGTATGGCAACCGCATGTGCCCTGCGCGCGAGA 238
|||
Db 558 CTGTGCTTTTATTAAGGTGGCGCTGTGTCCAGCGACATGGCGCACCGGCATTC--- 613
|||

Qy	239	CACAAACCCGAGCGAGCTTGGACACTGC-ACAGCGCATCTGGCCAGAGACTGTGACA	237
Db	614	-----CATCCAGCGAGCATCTGCCCGCCGCCGCCGCCCTCCAGAGACTGTGACA	669
Qy	298	CCACTTC 304	
Db	670	CCGCTCC 676	
RESULT 14			
XX	ABV75371	standard; DNA; 6128 BP.	
XX	ABV75371;		
XX	07-MAR-2003	(first entry)	
XX	Human IGFBP-1 gene sequence.		
XX	Insulin-like growth factor binding protein; IGFBP; cytotatic; liver;		
XX	cancer; human; IGFBP-1; gene; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
XX	CDS	762..5200	
XX		/*tag= a	
XX		/product= "IGFBP-1"	
XX		/note= "Insulin-like growth factor binding protein;	
XX		containing introns"	
XX	exon	762..1110	
XX		/*tag= b	
XX		/number= 1	
XX	intron	111..2656	
XX		/*tag= c	
XX		/number= 1	
XX	exon	2657..2826	
XX		/*tag= d	
XX		/number= 2	
XX	intron	2827..4040	
XX		/*tag= e	
XX		/number= 2	
XX	exon	4041..4169	
XX		/*tag= f	
XX		/number= 3	
XX	intron	4170..5068	
XX		/*tag= g	
XX		/number= 3	
XX	exon	5069..5197	
XX		/*tag= h	
XX		/number= 4	
XX	MO200290580-A1.		
XX	14-NOV-2002.		
XX	03-MAY-2002; 2002MO-AU000558.		
XX	03-MAY-2001; 2001US-0288441P.		
XX	(NACA-) NAT CANCER CENT SINGAPORE PTE LTD.		
XX	(SIGE-) SINGAPORE GEN HOSPITAL PTE LTD.		
XX	(ARGA/) ARGAET V P.		
XX	Huyh TH, Chow PKH, Soo KC;		
XX	WPI; 2003-103522/09.		
XX	P-PSDB; ABB82757.		
XX	Detecting the presence or diagnosing the risk of a liver cancer in a		
XX	patient comprises detecting aberrant expression of a gene encoding an		
XX	Insulin-like growth factor binding protein.		

Query Match	19.9%	Score 63.8	DB 10	Length 6128
Best Local Similarity	63.2%	Pred. No. 3.9e-09		
Matches 156	Conservative 0	Mismatches 72	Indels 19	Gaps 3
69	TCACGAAACACGGGAGTGC	CCCCGTGCGCCCATGTACAGCAAAACAACTTATTTTGA	128	
438	TCCCTCCACCGCGGCTTGGCTGTACGAGGCGCTTGAGCAAAACAACTTATTTTGA	497		
129	ACACGCGGGATCTTACGACGCGTGGCCCTGACATCATTTAAACC-----GTGCTGCG	178		
498	ACATCTACGCTCTTACGCGTGGCGCCCTGCAATCATTTAACTCTCTGTGCAAAAGTGGCGGG	557		
179	AGCAGCGCCTTCAATTAAGGCGCTGGGTATGCGCAGCCAGCATGTGTCACTGCGCCGCGAGA	238		
558	CGTGTGCGCTTTATTAAGTGTGGCGCTGTGTCTACAGAGCATCGGCGCACCGCATTC-----	613		
239	CACAAACCCAGCGAGCATTTGAACACTGC-ACACGCGCATCTTGCCCAAGAGCTGTACCA	297		
614	-----CATTCAGCGAGCATCTGCGCCGCGCGCGCGCACCATCTCCACAGAGACACTGGCCA	669		
298	CCACTTC	304		
670	CCGCTCC	676		

RESULT 15

AAH57489

AAH57489 standard; cDNA; 6134 BP.

AAH57489;

10-SEP-2001 (first entry)

Human liver cell specific cDNA sequence SEQ ID NO:329.

Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; sg; metabolic disease; developmental disease; cystostatic; immunomodulatory; neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

Homo sapiens.

W0200132927-A2.

10-MAY-2001.

02-NOV-2000; 2000MO-US030396.

04-NOV-1999; 99US-0163508P.

(INCY-) INCYTE GENOMICS INC.

Sornasse T, Sellhammer JJ, Watson GA;

WPI; 2001-291057/30.

New cell and tissue specific polynucleotides useful for diagnosis, prognosis or monitoring of treatments for disorders where the gene is associated with a cancer, immunopathology or neuropathology.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 12:52:03 ; Search time 394.464 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372

Sequence: 1 tacactcgtggggcgcagagtcgc.....ttccgctactagctagcgcg 372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn20018:*
5: geneeqn20028:*
6: geneeqn20038:*
7: geneeqn20048:*
8: geneeqn20058:*
9: geneeqn20068:*
10: geneeqn20078:*
11: geneeqn20088:*
12: geneeqn20098:*
13: geneeqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	372	100.0	372	6	ABX15375 Rat insul
2	321	86.3	321	6	ABX15374 Rat insul
3	223.6	60.1	423	6	ABX15376 Rat insul
4	219.4	59.0	270	6	ABX15373 Rat insul
5	219	58.9	219	6	ABX15372 Rat insul
6	154	41.4	423	6	ABX15376 Rat insul
7	76.4	20.5	1500	6	ABK63701 Rat
8	76.4	20.5	1500	10	ADBS58201
9	76.4	20.5	1500	10	ADBS52710
10	76.4	20.5	1500	10	ADBS52710
11	76.4	20.5	1500	12	ADP72689
12	76.4	20.5	1500	12	ADP72689
13	63.8	17.2	6128	6	ABN95896 Rat
14	63.8	17.2	6128	6	ABN95896 Rat
15	63.8	17.2	6128	6	ABN95896 Rat
16	55.2	14.8	13011	2	AAH57489 Human
17	55.2	14.8	13011	2	AAH57489 Human
18	55.2	14.8	13011	2	AAH57489 Human
19	55.2	14.8	13011	2	AAH57489 Human
20	51	13.7	51	6	ABX15380 Rat liver

ALIGNMENTS

21	51	13.7	51	6	ABX15371	Abx15371 Rat Gluco
22	51	13.7	270	6	ABX15373	Abx15373 Rat insul
23	48	12.9	48	6	ABX15379	Abx15379 Rat liver
24	38.6	10.4	4461	6	AA520000	AA520000 DNA encod
25	37.8	10.2	7061	6	ABL32997	ABL32997 Human imm
26	37.8	10.2	7061	6	ABL70248	ABL70248 Chemical1
27	37.8	10.2	7061	6	AA561189	AA561189 Human gen
28	37.6	10.1	1254	5	AA577024	AA577024 DNA encod
29	37.6	10.1	3191	12	ACH87111	ACH87111 Human gen
30	37.4	10.1	12850	10	ADBS6326	ADBS6326 Human fac
31	37	9.9	3896	12	ADQ64446	ADQ64446 Novel hum
32	36.6	9.9	761	6	ABL90266	ABL90266 Human pol
33	36.6	9.8	1678	10	ADA54241	ADA54241 Human cod
34	36.6	9.8	3054	6	AB570481	AB570481 Human bon
35	36.2	9.7	586	12	ACH71991	ACH71991 Human gen
36	35.8	9.6	11009	6	ABQ72907	ABQ72907 Mouse lam
37	35.8	9.6	11009	6	AA170816	AA170816 Mouse lam
38	35.6	9.6	490	9	ACH34858	ACH34858 Human end
39	35	9.4	3895	6	ABN87725	ABN87725 Human pro
40	34	9.1	10537	4	AA535769	AA535769 Human car
41	34	9.1	10537	4	AAK69582	AAK69582 Human imm
42	34	9.1	10537	10	AD546463	AD546463 Human car
43	34	9.1	10537	13	ADJ07881	ADJ07881 Human car
44	34	9.1	10543	4	AA535770	AA535770 Human car
45	34	9.1	10543	4	AAK69583	AAK69583 Human imm

RESULT 1

ABX15375
ID ABX15375 standard; DNA, 372 BP.

ABX15375;

17-APR-2003 (first entry)

Rat insulin regulator construct DNA #3.

Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGF1P-1; insulin-sensitive element; ISR; basal promoter; hyperglycaemia; insulin; insulin-like growth factor binding protein-1; hypoglycaemia; glucose; fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat; glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis; hepatocyte; hepatoma; cellular protein degradation; antidiabetic; intracellular hormone receptor; insulin regulator construct; anabolic.

Rattus norvegicus.

Synthetic.

US2002107198-A1.

08-AUG-2002.

10-OCT-2001; 2001US-00972916.

11-OCT-2000; 2000US-0239113P.

(THUL/) THUL P M.

Thule PM;

WPI; 2002-674190/72.

New insulin regulator cassette, useful e.g. for treating diabetes, provides specific, glucose-inducible transgenic expression of insulin in liver cells.

Claim 9; Page 14; 37p; English.

The invention relates to an insulin regulator construct comprising a glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 372; DB 6; Length 372;

Best Local Similarity 100.0%; Pred. No. 6.8e-103; Indels 0; Gaps 0;

Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACACTGGGGGCCAGAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGG 60
 DB 1 TACACTGGGGGCCAGAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGG 60
 QY 61 GGCACAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGGGCCAGAGT 120
 DB 61 GGCACAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGGGCCAGAGT 120
 QY 121 CCAGGAACCAAGGAGTCCCGGCGCCATGTACAAAGCAAAACCAATTATTTGAA 180
 DB 121 CCAGGAACCAAGGAGTCCCGGCGCCATGTACAAAGCAAAACCAATTATTTGAA 180
 QY 181 CACGGGAGTCTAGACAGCTGCTGACATATTAACCCGCTGCTGCCAGCAGCCCTT 240
 DB 181 CACGGGAGTCTAGACAGCTGCTGACATATTAACCCGCTGCTGCCAGCAGCCCTT 240
 QY 241 CATTAAGGCTCTGGGATATGCGCCAGCAGATGCTCCATCTGCGCGAGCAAAACCCAG 300
 DB 241 CATTAAGGCTCTGGGATATGCGCCAGCAGATGCTCCATCTGCGCGAGCAAAACCCAG 300
 QY 301 CGAGCATTTGAACACTGACACAGGCCATCTGCCAGAGAGCTGTGACCAACTTCGCGTA 360
 DB 301 CGAGCATTTGAACACTGACACAGGCCATCTGCCAGAGAGCTGTGACCAACTTCGCGTA 360
 QY 361 CTAGCTAGCCGC 372
 DB 361 CTAGCTAGCCGC 372

RESULT 2

ABX15374 standard; DNA; 321 BP.

AC ABX15374;

DT 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #2.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoids; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
 SS Synthetic.

XX US2002107198-A1.
 PN
 XX 08-AUG-2002.
 PD
 XX 10-OCT-2001; 2001US-00972916.
 PF
 XX 11-OCT-2000; 2000US-0239113P.
 PR
 XX (THUL/) THULE P M.
 PA
 XX Thule PM;
 PI
 XX WPI; 2002-674190/72.
 DR
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 PT
 XX
 XX
 XX
 PS Claim 9; Page 14; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 86.3%; Score 321; DB 6; Length 321;

Best Local Similarity 100.0%; Pred. No. 2e-87;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TACACTGGGGGCCAGAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGG 111
 DB 1 TACACTGGGGGCCAGAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACACTGGG 60
 QY 112 GGCACAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACAAAGCAAAACCAACT 171
 DB 61 GGCACAGTCCAGAAACCAAGGAGTCCCGGCGCCATGTACAAAGCAAAACCAACT 120
 QY 172 TATTTGAACACGGGGATCTAGACAGCTGCTGACATATTAACCCGCTGCGGAG 231
 DB 121 TATTTGAACACGGGGATCTAGACAGCTGCTGACATATTAACCCGCTGCGGAG 180
 QY 232 CCAGCCCTTCAATAGGCCCTGGTATGCGCCAGCAGAGATGCTCACTGCCCGCGAGACA 291
 DB 181 CCAGCCCTTCAATAGGCCCTGGTATGCGCCAGCAGAGATGCTCACTGCCCGCGAGACA 240
 QY 292 CAAACCCAGCAGATTTGAACACTGACACAGGCCATCTGCCAAGAGCTGTGACCA 351
 DB 241 CAAACCCAGCAGATTTGAACACTGACACAGGCCATCTGCCAAGAGCTGTGACCA 300
 QY 352 CTTCGGCTACTAGCTAGCCGC 372
 DB 301 CTTCGGCTACTAGCTAGCCGC 321

RESULT 3

ABX15376
ID ABX15376 standard; DNA; 423 BP.
XX
AC ABX15376;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #4.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
XX
Query Match 60.1%; Score 223.6; DB 6; Length 423;
Best Local Similarity 76.2%; Pred. No. 8,8e-58;
Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

QY 122 CAGGAACCAAGGAGTGGCCCGTGGCCCATG-----TCACAAACA 162
DB 154 CATGGGCGCAGCGGGGACACTCCGCTGGTTCCTGGACTGTGGCCCCAGGTGTATCACAAACA 213
QY 163 AAACAACTTATTTTGAACAAGGGATCTTACAGACGTGCGCTGACAAATCATTTAACCCGT 222
DB 214 AAACAACTTATTTTGAACAAGGGATCTTACAGACGTGCGCTGACAAATCATTTAACCCGT 273
QY 223 GCTGCGAGCCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCACGTGCC 282
DB 274 GTGCGCAGCCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCACGTGCC 333
QY 283 GCCGAGACCAAAACCCAGCAGCATTTGAACACTGCACACGCGCATCTGCCCCAGAGACTG 342
DB 334 GCCGAGACCAAAACCCAGCAGCATTTGAACACTGCACACGCGCATCTGCCCCAGAGACTG 393
QY 343 TGACCAACCACTTCGCTACTAGCTAGCCGC 372
DB 394 TGACCAACCACTTCGCTACTAGCTAGCCGC 423
RESULT 4
ABX15373
ID ABX15373 standard; DNA; 270 BP.
XX
AC ABX15373;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 13-14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis.
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 59.0%; Score 219.4; DB 6; Length 270;

Best Local Similarity 99.5%; Pred. No. 1.4e-56;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 152 TGTCAAGCAAAACAACTTATTTTGAACACGGGGATCTTACACGCTGCCCTTACAT 211
DB 50 TATCAAGCAAAACAACTTATTTTGAACACGGGGATCTTACACGCTGCCCTTACAT 109
OY 212 CATTAACCGGCTGCGGAGCAGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGCATG 271
DB 110 CATTAACCGGCTGCGGAGCAGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGCATG 169
OY 272 GTTCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACATCTCACAGGCCATCTGC 331
DB 170 GTTCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACATCTCACAGGCCATCTGC 229
OY 332 CCAAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCGC 372
DB 230 CCAAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCGC 270

RESULT 5

ABX15372
ID ABX15372 standard; DNA; 219 BP.

AC ABX15372;

DT 17-APR-2003 (first entry)

DE Rat insulin-sensitive element (ISE) DNA.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

PA (THUL/) THULE P M.

PI Thule PM;

DR WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

PS Claim 8; Page 13; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention
XX

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 58.9%; Score 219; DB 6; Length 219;

Best Local Similarity 100.0%; Pred. No. 1.8e-56;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 TCACAGCAAAACAACTTATTTGAACACGGGGATCTTACACGCTGCCCTTACATCA 213
DB 1 TCACAGCAAAACAACTTATTTTGAACACGGGGATCTTACACGCTGCCCTTACATCA 60
OY 214 TTAACCCGCTGCTGCGGAGCAGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGCATG 273
DB 61 TTAACCCGCTGCTGCGGAGCAGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGCATG 120
OY 274 CCACTGCCCGCGGAGACAAACCCAGCAGCATTTGAACATCTGACACGGCCATCTGCC 333
DB 121 CCACTGCCCGCGGAGACAAACCCAGCAGCATTTGAACATCTGACACGGCCATCTGCC 180
OY 334 AGAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCGC 372
DB 181 AGAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCGC 219

RESULT 6

ABX15376/C
ID ABX15376 standard; DNA; 423 BP.

AC ABX15376;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #4.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN Synthetic.

PD US2002107198-A1.

PF 08-AUG-2002.

PR 10-OCT-2001; 2001US-00972916.

PA (THUL/) THULE P M.

```

XX Thule PW;
PI MPI; 2002-674190/72.
DR MPI; 2002-674190/72.
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX Claim 9; Page 14, 37pp; English.
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
XX
SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
Query Match 41.4%; Score 154; DB 6; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.2e-36;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TACACTGGGGGGCCAGAGTCCAGAAACCAACGGAGTCCCGCCGCCCATGTACACTGGG 60
DB 204 TACACTGGGGGGCCAGAGTCCAGAAACCAACGGAGTCCCGCCGCCCATGTACACTGGG 145
QY 61 GGCAGAGTCCAGAAACCAACGGAGTCCCGCCGCCCATGTACACTGGGGCCAGAGT 120
DB 144 GGCAGAGTCCAGAAACCAACGGAGTCCCGCCGCCCATGTACACTGGGGCCAGAGT 85
QY 121 CCAGAACCCAGCGAGTGGCCCGCCGCCCATGT 154
DB 84 CCAGAACCCAGCGAGTGGCCCGCCGCCCATGT 51
RESULT 7
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
XX
XX ABK63701;
AC 18-JUN-2002 (first entry)
XX
XX 18-JUN-2002 (first entry)
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX Rat; as; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
OS
XX
XX WO200210453-A2.
XX
XX 07-FEB-2002.
XX
XX 30-JUL-2001; 2001WO-US023872.
XX
XX 31-JUL-2000; 2000US-0222040P.
XX
XX 02-NOV-2000; 2000US-0244880P.

```

```

PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX MPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
XX method can also be used to identify an agent which modulates the toxic
XX response and predict cellular pathways that a compound modulates in a
XX cell. The method utilizes a set of at least two probes (on a solid
XX support in kit form), where each of the probes comprises a sequence that
XX specifically hybridises to a gene listed in the specification, a computer
XX system comprising a database containing information identifying the
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX set of genes comprising at least two genes listed in the specification,
XX and a user interface to view the information used to present information,
XX CC identifying the expression level in a tissue or cell of at least one gene
XX listed in the specification. The method is useful for elucidating global
XX changes in gene expression and for identifying toxicity markers in
XX tissues or cell exposed to a known toxin. The genes may be used as
XX toxicity markers in drug screening and toxicity assays. The genes and
XX gene expression information may be used as diagnostic markers for the
XX prediction or identification of the physiological state of tissue or cell
XX sample that has been exposed to a compound or agent. Hepatotoxicity is
XX characterised by centrilobular necrosis and steatosis. The present
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX which is differentially expressed in response to a hepatotoxic agent
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13; 1; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAACCCAGCGAGTTCAGTCACTGCACAGCGCCATCTCCAGAGAGCTGTACCAC 349
DB 8 CACAACCCAGCGAGTTCAGTCACTGCACAGCGCCATCTCCAGAGAGCTGTACCAC 67
QY 350 CACTTCGCTACTAGCTA 367
DB 68 CACTTCGCTACTATCTA 85
RESULT 8
ADB58201
ID ADB58201 standard; DNA; 1500 BP.
XX
XX ADB58201;
AC

```

DT 04-DEC-2003 (first entry)
XX Toxicity-related gene, SEQ ID 3227.
DE
XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KM drug screening; toxicity assay; ds.
XX
XX Unidentified.
OS
XX WO2003064624-A2.
PN
XX
XX 07-AUG-2003.
PD
XX
XX 31-JAN-2003; 2003WO-US003194.
PF
XX
PR 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M,
PI WPI; 2003-689530/65.
DR
XX
XX Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
the compound.
XX
XX Claim 1; SEQ ID NO 3227; 1156bp; English.
XX
XX The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAAACCCGAGCATTTGAACACTGACACGGCCATCTGCCGAGAGCTGTGACAC 349
DB 8 CACAAACCCGAGCATTTGAACACTGACACGGCCATCTGCCGAGAGCTGTGACAC 67
QY 350 CACTTCCGCTACTAGCTA 367
DB 68 CACTTCCGCTACTAGCTA 85
RESULT 9
ADBS2710
ID ADBS2710 standard; DNA; 1500 BP.
XX
AC ADBS2710;
XX
DT 04-DEC-2003 (first entry)

XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX Rattus norvegicus.
OS
XX WO2003065993-A2.
PN
XX
XX 14-AUG-2003.
PD
XX
XX 04-FEB-2003; 2003WO-US003482.
PF
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M,
PI Elashoff M;
XX WPI; 2003-731472/69.
DR
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3252; 874bp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAAACCCGAGCATTTGAACACTGACACGGCCATCTGCCGAGAGCTGTGACAC 349
DB 8 CACAAACCCGAGCATTTGAACACTGACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 350 CACTTCGCGTACTAGCTA 367
|||||
68 CACTTCGCGTACTATCTA 85
Db

RESULT 10
ABT41911
ID ABT41911 standard; DNA; 1500 BP.
XX
XX ABT41911;
XX
XX 26-JUN-2003 (first entry)
XX
XX Toxicity modelling related rat gene SEQ ID NO 1613.
DE
XX Toxic effect; gene expression profile; renal toxicity; toxicity marker;
XX database; drug screening; toxicity assay; rat; ds.
KM
XX Rattus norvegicus.
OS
XX WO200295000-A2.
PN
XX 28-NOV-2002.
PD
XX 22-MAY-2002; 2002WO-US016173.
PF
XX
XX 22-MAY-2001; 2001US-0299233P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-SEP-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI
XX WPI; 2003-148464/14.
DR
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modelling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX Example 4; Page; 446pp; English.
PS
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the renal
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide

CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
XX SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 230 CACAACCAGGAGCATTTGAACACTGCACGCGCATCTGCGGAGAGCTGTACAC 349
|||||
Db 8 CACAACCAGGAGCATTTGAACACTGCACGCGCATCTGCGGAGAGCTGTACAC 67
QY 350 CACTTCGCGTACTAGCTA 367
|||||
Db 68 CACTTCGCGTACTATCTA 85

RESULT 11
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
XX ADP72689;
XX
XX 26-AUG-2004 (first entry)
XX
XX Renal toxin progression gene marker #1278.
DE
XX ds; toxic effect; gene expression profile; kidney tissue;
KM differential gene expression; toxicity progression; toxicity marker;
KM drug screening; toxicity assay; kidney pathology; nephritis;
KM kidney necrosis; glomerular injury; tubular injury;
KM focal segmental glomerulosclerosis.
XX
XX Rattus norvegicus.
OS
XX WO2004048598-A2.
PN
XX 10-JUN-2004.
PD
XX 24-NOV-2003; 2003WO-US037556.
PF
XX 22-NOV-2002; 2002US-00301856.
PR
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B,
PI Elashoff M;
PI
XX WPI; 2004-460771/43.
DR
XX
XX Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
XX Claim 11; SEQ ID NO 1278; 266pp; English.
PS
XX The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. Predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC method are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental

XX 14-NOV-2002.
PD
XX
XX
PF 03-MAY-2002; 2002WO-AU000558.
XX
XX
PR 03-MAY-2001; 2001US-0288441P.
XX
XX
PA (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
PA (SIGE-) SINGAPORE GEN HOSPITAL PTE LTD.
PA (ARGA/) ARGAET V P.
XX
XX
PI Huiyuh TH, Chow PKH, Soo KC;
XX
XX
DR WPI; 2003-1035522/09.
DR P-PSDB; ABB82757.
XX
PT Detecting the presence or diagnosing the risk of a liver cancer in a
PT patient comprises detecting aberrant expression of a gene encoding an
PT insulin-like growth factor binding protein.
XX
PS
XX Example; Page 104-108; 142pp; English.

XX The invention relates to detecting the presence or diagnosing the risk of
CC a liver cancer in a patient. The method involves detecting in a
CC biological sample obtained from the patient aberrant expression of a gene
CC encoding an insulin-like growth factor binding protein (IGFBP). The
CC method is useful for detecting the presence or diagnosing the risk of a

XX	Sequence	6128 BP, 1504 A, 1497 C, 1504 G, 1623 T, 0 U, 0 Other;
CC	Best Local Similarity	17.2%; Score 63.8; DB 10; Length 6128;
CC	Match	63.2%; Pred. No. 7,3e-09;
CC	Matches	156; Conservative 0; Mismatches 72; Indels 19; Gaps 3
XX	encoding genomic DNA (Genbank Accession No. M74587)	
XX	Sequence	6128 BP, 1504 A, 1497 C, 1504 G, 1623 T, 0 U, 0 Other;
QY	120	TCACAGAAACACGGAGATGCCCCCTGTGCGCCCATGTACACAGAAACAACTATTATTGGA 179
Db	438	TCCCTCCACACAGCGATTGCGCTAGAGGCTTGGGTGCACATGACAAACTTATTTTTGA 497
QY	180	ACACGGGAGATCTTAGACAGCTGCCCTGCACATCATTAATCC-----GTGCTGCCG 229
Db	498	ACACTCAGCTCTTAGCGCTGCCGCTGCCCAATCATTAACCTCTGTGCATGATGGCCGG 557
QY	230	AGCAGCCCTTCATTAAGGCGCTGGGTATAGCCACGACAGCATGTGTCACCTGCCGCCGAGA 289
Db	558	CCTGTGCCCTTTATTAAGGTGCGCCCTGTGTGCAGGAGACATGTGGCCACCGCCATCC---- 613
QY	290	CACAAACCCACGAGCATTTGAACACTGC-ACACGGCCCATCTGCCAGAGAGCTGTGACCA 348
Db	614	----CATCCACGAGCATCTGTCCGCGCGCCGCCGCCACCTCCAGAGAGCATCTGGCCA 669
QY	349	CCACTTC 355
Db	670	CCGCTCC 676
XX	RESULT 15	
XX	AAH57489	
ID	AAH57489	standard; cDNA, 6134 BP.
XX	AAH57489;	
XX	10-SEP-2001	(first entry)
XX	Human Liver cell specific cDNA sequence SEQ ID NO:329.	
XX	Human, tissue specific; diagnosis; brain; heart; skeletal muscle; lung;	
KW	Liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;	
KW	metabolic disease; developmental disease; cytosolic; immunomodulatory;	
KW	neuroprotective; gene therapy; cancer; immunopathology; neuropathology.	

```
XX Homo sapiens.
OS
XX MO200132927-A2.
XX
XX 10-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-US030396.
XX
XX 04-NOV-1999; 99US-0163508P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Sornasse T, Seilhamer JJ, Watson GA;
XX
XX WPI; 2001-291057/30.
XX
XX New cell and tissue specific polynucleotides useful for diagnosis,
XX prognosis or monitoring of treatments for disorders where the gene is
XX associated with a cancer, immunopathology or neuropathology.
XX
XX Claim 1, Page 246-248; 327pp; English.
XX
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
XX sequences (I). (I) can have cytosstatic, immunomodulatory and
XX neuroprotective activities, and can be used in gene therapy. (I) and
XX proteins (II) encoded by then are used in high throughput screening
XX assays to select DNA molecules, RNA molecules, peptide nucleic acids,
XX mimetics, peptides, proteins, agonists, antagonists, antibodies or their
XX fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
XX agents. Expression of (I) in a sample indicates the differentiation of
XX embryonic stem cells into a tissue selected from brain, heart, kidney,
XX liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
XX to produce an expression profile that defines a metabolic or
XX developmental process, treatment, condition, disease or disorder. The
XX gene profile can be used for diagnosis, prognosis or monitoring of
XX treatments and for investigating a predisposition to a disorder where the
XX gene is associated with a cancer, immunopathology or neuropathology
XX
XX
SQ Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;
Query Match 17.2%; Score 63.8; DB 4; Length 6134;
Best Local Similarity 63.2%; Pred. No. 7.3e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
QY 120 TCCAGAGACCAAGGAGTCCCGCCGATGTCACAGCAAACTATTGGA 179
DB 438 TCCTCCACACGCGTTTGCGTAGGACCTGGGTGCACTAGCAAACTATTGGA 497
QY 180 ACACGGGATCCTAGCAGCTGCGCTGACATCATTAACCC-----GTGCTGCCG 229
DB 498 ACACCTCAGTCTTACGTCGCGCGCTGCCAATCATTAACCTCTGTGCAAGTGGCGGG 557
QY 230 AGCCAGCCCTTCATTAAGGCGCTGTGATGCGCAGCAGCATGTCCACTGCCCGCGAGA 289
DB 558 CCTGTGCCCTTTATTAAGTGCAGCTGTGTCCAGGAGATGAGCCACCGCATCC--- 613
QY 290 CACAAACCCAGGAGCATTTGAACATGCG-ACACGGCCATCTGCCAGAAAGCTGTGACCA 348
DB 614 ---CATCAGAGGACATCTGCGCGCGCGCGCCACCCCTCCAGAGAGCATGGCCA 669
QY 349 CCACTTC 355
DB 670 CCGCTCC 676
```

Search completed: September 1, 2005, 00:34:47
Job time : 396.464 secs

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 423; DB 6; Length 423;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGACGCGGCACTCCGCTGCTCTGACTCTGAGCCGCCAGTGTATAGGCGCG 60
DB 1 CATGGGCGACGCGGCGCACTCCGCTGCTCTGACTCTGAGCCGCCAGTGTATAGGCGCG 60
QY 61 ACGGGGCACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGAC 120
DB 61 ACGGGGCACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGAC 120
QY 121 TCCCGGTGTTCTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGACTCCGCTGT 180
DB 121 TCCCGGTGTTCTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGACTCCGCTGT 180
QY 181 TCTGTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGCGCGCGCGCGCGCGCGAT 240
DB 181 TCTGTGACTCTGAGCCGCCAGTGTATAGGCGCGCGCGCGCGCGCGCGCGCGCGAT 240
QY 241 CCTAGACGCTGCGCCCGTGTATAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 CCTAGACGCTGCGCCCGTGTATAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CTGGGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CTGGGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 AACACTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 AACACTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 CGC 423
DB 421 CGC 423

RESULT 2

ABX15373

ID ABX15373 standard; DNA; 270 BP.

AC ABX15373;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #1.

XX Glucose response element; GIRE; liver pyruvate kinase; lpk; de; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;

KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS Synthetic.
PN US2002107198-A1.
PD 08-AUG-2002.
PF 10-OCT-2001; 2001US-00972916.
PR 11-OCT-2000; 2000US-0239113P.
PA (THULE/) THULE P M.
PI Thule PM;
XX WPI; 2002-674190/72.
DR
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 13-14; 37pp; English.

XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match Best Local Similarity 63.8%; Score 270; DB 6; Length 270;

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 CATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 213
DB 1 CATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 214 AAACAACTTATTTTGAACAGCGGCGATCTGACAGCGTGCCTGACATCTTAACCGGT 273
DB 61 AAACAACTTATTTTGAACAGCGGCGATCTGACAGCGTGCCTGACATCTTAACCGGT 120
QY 274 GCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 333
DB 121 GCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 334 GCCGAGACAAACCCGAGCGCATTTGAACACTGACACGCGCGCATTTGCCAGAGAGCTG 393
DB 181 GCCGAGACAAACCCGAGCGCATTTGAACACTGACACGCGCGCATTTGCCAGAGAGCTG 240
QY 394 TGACACACCTTCGCTACTAGTAGCGCGC 423
DB 241 TGACACACCTTCGCTACTAGTAGCGCGC 270

RESULT 3

ABX15375
 ID ABX15375 standard; DNA; 372 BP.
 AC ABX15375;
 DT 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #3.
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase, LPK; db; IGFBP-1;
 KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX
 PN US2002107198-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THUL P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SO Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
 XX
 Query Match 52.9%; Score 223.6; DB 6; Length 372;
 Best Local Similarity 76.2%; Pred. No. 4.3e-59;
 Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;
 XX
 QY 34 ACTCTGGGCCCCAGGTATACATGAGGCGCACTCCCGTTCCTGAGCTCTGGGC 93
 DB 2 AACTCTGGGCCCAAGTCCAGAAACACAGGAGTGCCTCCGCGCCCACTGACACTGGGG 61
 QY 94 CCCAGTGTATCATGAGGCGCACTCCCGTGTCTCTGAGACTCTGAGCCGCCAGTGT 153
 DB 62 GCCAGAGTCCAGGAACACAGGAGTGCCTCCGCGCCCACTGATCACTGAGGCGCCAGAGTC 121

QY 154 CATGGCGCGACGGGCACTCCGTGTCTCTGAGCTTGAGCCCCCAGTATACAAAGCA 213
 DB 122 CAGAAACCAAGGAGTGCCTCCGTGCGCCCATG-----TCACAAAGCA 162
 QY 214 AAACAACTTATTTTGAACACGGGAGATCTAGACAGCTGCGCTGACATATTAAACCGGT 273
 DB 163 AAACAACTTATTTTGAACACGGGAGATCTAGACAGCTGCGCTGACATATTAAACCGGT 222
 QY 274 GCTGCGAGCCAGCCCTTCATTAAGCCCTGAGTATGCGCAGCAGCATGTCACATGCCCC 333
 DB 223 GCTGCGAGCCAGCCCTTCATTAAGCCCTGAGTATGCGCAGCAGCATGTCACATGCCCC 282
 QY 334 GCCGAGACAAACCCAGGAGATTTGAACATCTGACACAGGCACTCTGCCAGAGACTG 393
 DB 283 GCCGAGACAAACCCAGGAGATTTGAACATCTGACACAGGCACTCTGCCAGAGACTG 342
 QY 394 TGACCAACCACTTCGCTACTAGCTAGCCGC 423
 DB 343 TGACCAACCACTTCGCTACTAGCTAGCCGC 372
 XX
 RESULT 4
 ID ABX15374 standard; DNA; 321 BP.
 AC ABX15374;
 DT 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #2.
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase, LPK; db; IGFBP-1;
 KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX
 PN US2002107198-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THUL P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX

SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 51.9%; Score 219.4; DB 6; Length 321;
Best Local Similarity 99.5%; Pred. No. 8.2e-58;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 TATCACAAGCAAACTTATTTTGAACAGGGGATCTTGACAGCTGCTGACAT 262
DB 101 TGTCAACAGCAAACTTATTTTGAACAGGGGATCTTGACAGCTGCTGACAT 160
QY 263 CATTAAACCGTGCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGATAGCCAGCAGCATG 322
DB 161 CATTAAACCGTGCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGATAGCCAGCAGCATG 220
QY 323 GTCCACTGCGCCGCGAGACACAACCCAGCAGCATTTGAACACTGCACAGGCCATCTGC 382
DB 221 GTCCACTGCGCCGCGAGACACAACCCAGCAGCATTTGAACACTGCACAGGCCATCTGC 280
QY 383 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCCGC 423
DB 281 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCCGC 321

RESULT 5

ABX15372
ID ABX15372 standard; DNA; 219 BP.

AC ABX15372;

DT 17-APR-2003 (first entry)

DE Rat insulin-sensitive element (ISE) DNA.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THULE P M.

XX Thule PM;

XX WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

PS Claim 8; Page 13; 37BP; English.

XX

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention
XX

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 51.8%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 9.5e-58;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 TCACAGCAAAACAACTTATTTGAACAGGGGATCTTGACAGCTGCTGACATCA 264
DB 1 TCACAGCAAAACAACTTATTTGAACAGGGGATCTTGACAGCTGCTGACATCA 60
QY 265 TTAACCGTGCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGATAGCCAGCAGCATG 324
DB 61 TTAACCGTGCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGATAGCCAGCAGCATG 120
QY 325 CCACTGCGCCGCGAGACACAACCCAGCAGCATTTGAACACTGCACAGGCCATCTGCC 384
DB 121 CCACTGCGCCGCGAGACACAACCCAGCAGCATTTGAACACTGCACAGGCCATCTGCC 180
QY 385 AGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCCGC 423
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCCGC 219

RESULT 6

ABX15375/C
ID ABX15375 standard; DNA; 372 BP.

AC ABX15375;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #3.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THULE P M.

XX

XX Thule PM;
 PI MPI; 2002-674190/72.
 DR
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 PS
 XX Claim 9; Page 14; 37pp; English.
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetes during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
 Query Match 36.4%; Score 154; DB 6; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.6e-37;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 51 ACATGGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCG 110
 Db 154 ACATGGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCG 95
 Oy 111 CACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCGCACGGGCGA 170
 Db 94 CACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCGCACGGGCGA 35
 Oy 171 CTCCTGGTCTCTGACTCTGGCCCCAGTGT 204
 Db 34 CTCCTGGTCTCTGACTCTGGCCCCAGTGT 1
 RESULT 7
 ABX15374/c
 ID ABX15374 standard; DNA; 321 BP.
 AC
 XX ABX15374;
 XX
 DT 17-APR-2003 (first entry)
 XX
 DE Rat insulin regulator construct DNA #2.
 XX
 XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 OS
 XX US2002107198-A1.
 XX

PD 08-AUG-2002.
 XX
 XX 10-OCT-2001; 2001US-00972916.
 XX
 XX 11-OCT-2000; 2000US-0239113P.
 XX
 XX (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR MPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 PS
 XX Claim 9; Page 14; 37pp; English.
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetes during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
 Query Match 24.3%; Score 103; DB 6; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 51 ACATGGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCG 110
 Db 103 ACATGGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTACATGGGCG 44
 Oy 111 CACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGT 153
 Db 43 CACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGT 1
 RESULT 8
 ABK63701
 ID ABK63701 standard; cDNA; 1500 BP.
 AC
 XX ABK63701;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatocxin #1608.
 KW Rat; ss; hepatocxin; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 XX
 OS Rattus norvegicus.
 OS
 XX WO200210453-A2.
 XX
 XX 07-FEB-2002.
 PD
 XX 30-JUL-2001; 2001WO-US023872.
 PF

```

XX 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
XX WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
XX method can also be used to identify an agent which modulates the toxic
XX response and predict cellular pathways that a compound modulates in a
XX cell. The methods utilize a set of at least two probes (on a solid
XX support in kit form), where each of the probes comprises a sequence that
XX specifically hybridises to a gene listed in the specification, a computer
XX system comprising a database containing information identifying the
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX set of genes comprising at least two genes listed in the specification,
XX and a user interface to view the information used to present information
XX identifying the expression level in a tissue or cell of at least one gene
XX listed in the specification. The method is useful for elucidating global
XX changes in gene expression and for identifying toxicity markers in
XX tissues or cell exposed to a known toxin. The genes may be used as
XX toxicity markers in drug screening and toxicity assays. The genes and
XX gene expression information may be used as diagnostic markers for the
XX prediction or identification of the physiological state of tissue or cell
XX sample that has been exposed to a compound or agent. Hepatotoxicity is
XX characterised by centrilobular necrosis and steatosis. The present
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX which is differentially expressed in response to a hepatotoxic agent
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 6; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 341 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 400
DB 8 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
QY 401 CACTTCGCGTACTAGCTA 418
DB 68 CACTTCGCGTACTAGCTA 85

```

```

RESULT 9
ADBS8201
ID ADBS8201 standard; DNA; 1500 BP.

```

```

XX ADBS8201;
XX
XX 04-DEC-2003 (first entry)
XX
XX Toxicity-related gene, SEQ ID 3227.
XX
XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
XX drug screening; toxicity assay; ds.
XX
XX Unidentified.
XX
XX WO2003064624-A2.
XX
XX 07-AUG-2003.
XX
XX 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-00060087.
XX 15-MAR-2002; 2002US-0364045P.
XX 15-MAR-2002; 2002US-0364055P.
XX 30-DEC-2002; 2002US-0436643P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-689530/65.
XX
XX Predicting a toxic effect of a compound, useful in identifying toxicity
XX markers in liver tissues or cells for drug screening and toxicity assays,
XX comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.
XX
XX Claim 1; SEQ ID NO 3227; 1156pp; English.
XX
XX The present invention relates to a method for predicting a toxic effect
XX of a compound. The method comprises preparing a gene expression profile
XX of a tissue or cell sample exposed to the compound, and comparing the
XX gene expression profile to a database comprising SEQ ID 1-4925, where
XX differential expression of the gene indicates at least one toxic effect.
XX The method is useful for predicting at least one toxic effect of a
XX compound, predicting hepatotoxicity or the progression of a toxic effect
XX of a compound, identifying an agent that modulates the onset or
XX progression of a toxic response, predicting the cellular pathways that a
XX compound modulates in a cell, and identifying an agent that modulates at
XX least one activity of a protein. The method and compositions of the
XX present invention using a database of genes having liver toxin-induced
XX differential expression, are useful in identifying toxicity markers in
XX liver tissues or cells for drug screening and toxicity assays. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 10; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 341 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 400
DB 8 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
QY 401 CACTTCGCGTACTAGCTA 418
DB 68 CACTTCGCGTACTAGCTA 85

```

```

RESULT 10
ADBS2710
ID ADBS2710 standard; DNA; 1500 BP.

```

AC ADB52710;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
KM toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/59.
DR
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3252; 874pp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 18.1%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.1e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

341 CACAAACCCAGCAGCATTTGACACACGCGCCATCTGCCGAGAGCTGTACAC 400

DB 8 CACAAACCCAGCAGCATTTGACACACGCGCCATCTGCCGAGAGCTGTACAC 67
QY 401 CACTTCGCTACTACTA 418
DB 68 CACTTCGCTACTACTA 85
RESULT 11
ID ABT41911
ID ABT41911 strand; DNA; 1500 BP.
XX
AC ABT41911;
XX
DT 26-JUN-2003 (first entry)
XX
DE Toxicity modelling related rat gene SEQ ID No 1613.
XX
KM Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KM database; drug screening; toxicity assay; rat; ds.
XX
OS Rattus norvegicus.
XX
PN WO200295000-A2.
XX
PD 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
PR 22-MAY-2001; 2001US-0292335P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI WPI; 2003-148464/14.
DR
PT Predicting at least one toxic effect of a compound, useful for toxicity
PT modelling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
PS Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
XX effect of a compound. The method comprises a gene expression profile of a
XX tissue or cell sample exposed to the compound, and comparing the gene
XX expression profile to a database comprising at least part of the data or
XX information given in the specification. The methods are useful for
XX predicting at least one toxic effect of a compound, predicting the
XX progression of a toxic effect of a compound, predicting the renal
XX toxicity of a compound, or identifying toxicity markers in tissues or

CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization

XX SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 10; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCGAGCATTGAACACTGCACGCGCATCTGCCGAGAGCTGTGACAC 400
DB 8 CACAAACCCGAGCATTGAACACTGCACGCGCATCTGCCGAGAGCTGTGACAC 67
QY 401 CACTTCGCTACTAGCTA 418
DB 68 CACTTCGCTACTATCTA 85

RESULT 12
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
XX ADP72689;
XX
XX 26-AUG-2004 (first entry)
XX
XX
XX Renal toxin progression gene marker #1278.
XX
XX
XX ds; toxic effect; gene expression profile; kidney tissue;
XX differential gene expression; toxicity progression; toxicity marker;
XX drug screening; toxicity assay; kidney pathology; nephritis;
XX kidney necrosis; glomerular injury; tubular injury;
XX focal segmental glomerulosclerosis.
XX
XX Rattus norvegicus.
XX
XX WO2004048598-A2.
XX
XX 10-JUN-2004.
XX
XX 24-NOV-2003; 2003WO-US037556.
XX
XX 22-NOV-2002; 2002US-00301856.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
XX Elashoff M;
XX
XX WPI; 2004-460771/43.
XX
XX Predicting (the progression of) a toxic effect of a compound, for
XX monitoring the progression of renal disease states, comprises preparing a
XX gene expression profile of a kidney tissue or cell sample exposed to the
XX compound.
XX
XX Claim 11; SEQ ID NO 1278; 266pp; English.

CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The

CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 12; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCGAGCATTGAACACTGCACGCGCATCTGCCGAGAGCTGTGACAC 400
DB 8 CACAAACCCGAGCATTGAACACTGCACGCGCATCTGCCGAGAGCTGTGACAC 67
QY 401 CACTTCGCTACTAGCTA 418
DB 68 CACTTCGCTACTATCTA 85

RESULT 13
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
XX
XX AAH22429;
XX
XX 22-AUG-2001 (first entry)
XX
XX
XX Rat insulin-like growth factor binding protein nucleotide sequence.
XX
XX
XX Identification; toxic; hepatotoxic; differential gene expression; NSAID;
XX non-steroidal antiinflammatory drug; ds.
XX
XX Rattus norvegicus.
XX
XX WO200138579-A2.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US032049.
XX
XX 22-NOV-1999; 99US-0166923P.
XX 18-FEB-2000; 2000US-0183531P.
XX 20-NOV-2000; 2000US-00717321.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gould-Rothberg BE, Dipippo VA, Ramseh TM, Gerwein RW;
XX WPI; 2001-355948/37.
XX
XX Screening hepatotoxic agent comprises contacting test cell population
XX expressing RISKMARKER or INJURYMARKER with agent, comparing expression
XX with reference population and identifying difference in expression
XX levels.
XX
XX Disclosure; Page 22-24; 76pp; English.

CC The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test

CC agent; (c) measuring expression of one or more of the nucleic acid
 CC sequences in the test cell population; (d) comparing the expression of
 CC the nucleic acid sequence in the test cell population to the expression
 CC of the nucleic acid sequence in a reference cell population comprising
 CC at least one cell whose exposure status to a hepatotoxic agent is known;
 CC and (e) identifying a difference in expression levels of the RISKMARKER
 CC or INJURYMARKER sequences, if present, in the test cell population and
 CC reference cell population. The method is useful for identifying a
 CC hepatotoxic agent. The present sequence is given in the exemplification
 CC of the present invention

XX SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

XX Query Match 18.1%; Score 76.4; DB 4; Length 5001;
 XX Best Local Similarity 98.7%; Pred. No. 4.8e-13;
 XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCAGCAGCATTTGAACACTGACACGCGCCATCGCCAGAGAGCTGTGACAC 400
 DB 1 CACAAACCCAGCAGCATTTGAACACTGACACGCGCCATCGCCAGAGAGCTGTGACAC 60
 QY 401 CACTTCGCTACTAGCTA 418
 DB 61 CACTTCGCTACTAGCTA 78

XX RESULT 14
 XX ABN95896
 XX ID ABN95896 standard; DNA; 6128 BP.
 XX AC ABN95896;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #2394 used to diagnose liver cancer.
 XX KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 XX metastatic liver tumor; cytostatic; expression profile; disease state;
 XX disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX PN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US030589.
 XX PR 02-OCT-2000; 2000US-0237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvaras C, Peres-Da-Silva S, Vockley JG;
 XX DR WPI; 2002-426119/45.
 XX PT Diagnosing and detecting the progression of liver cancer, hepatocellular
 XX carcinoma or metastatic liver tumor in a patient, involves detecting the
 XX level of expression of two or more genes in a liver tissue sample.
 XX PS Claim 1, SEQ ID NO 2394; 298BP; English.

XX The invention relates to a novel method for diagnosing and detecting the
 XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
 XX tumor in a patient, and differentiating metastatic liver cancer from
 XX hepatocellular carcinoma in a patient, involving detecting the level of
 XX expression of two or more genes represented in ABN93503-ABN97455 in a
 XX tissue sample. The method of the invention has hepatotropic, and
 XX cytostatic activity. The method is useful for diagnosing and detecting
 XX the progression of liver cancer, hepatocellular carcinoma and metastatic
 XX liver carcinoma in a patient. The method is useful for identifying
 XX expression profiles which serve as useful diagnostic markers as well as
 XX markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WFO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;

XX Query Match 14.4%; Score 60.8; DB 6; Length 6128;
 XX Best Local Similarity 66.5%; Pred. No. 3.5e-08;
 XX Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTATTGTAACACGGGATCTAGACAGCTGCCATGATCAT 265
 DB 473 CACTAGCAAAACAACTATTGTAACACAGCTAGCTGCGCGCTGCCATCAT 532
 QY 266 TAACCC-----GTGCTGCGCAGCCAGCCCTTCATTAAGCCCTGGATAGCCAGC 315
 DB 533 TAACCTCTGTGTAAGTGGCGCGGCTGTGCTTTTAAGGTGCGGCTGTGCAAGC 592
 QY 316 CAGCATGTGTCACATGCGCGCGCGGAGACACAAACCCAGCGATTTGAACACTGC 374
 DB 593 GAGCATGCGCCACCGCCATTC-----CATCCGAGGAGCATTCGCCGCGCGCGCG 644
 QY 375 CCATCGCCAGAGAGCTGTGACACCACTTC 406
 DB 645 CCACCTCCAGAGGACACTGCGCACCGCTCC 676

XX RESULT 15
 XX ABV75371
 XX ID ABV75371 standard; DNA; 6128 BP.
 XX AC ABV75371;
 XX DT 07-MAR-2003 (first entry)
 XX DE Human IGFBP-1 gene sequence.
 XX KM Insulin-like growth factor binding protein; IGFBP; cytoablastic; liver;
 XX cancer; human; IGFBP-1; gene; ds.
 XX OS Homo sapiens.
 XX PN WO200290580-A1.
 XX PD Key CDS
 XX FT Location/Qualifiers
 XX FT 762..5200
 XX FT /tag= a
 XX FT /product= "IGFBP-1"
 XX FT /note= "insulin-like growth factor binding protein;
 XX contains introns"
 XX FT 762..1110
 XX FT /tag= b
 XX FT /number= 1
 XX FT 1111..2656
 XX FT /tag= c
 XX FT /number= 1
 XX FT 2657..2826
 XX FT /tag= d
 XX FT /number= 2
 XX FT 2827..4040
 XX FT /tag= e
 XX FT /number= 2
 XX FT 4041..4169
 XX FT /tag= f
 XX FT /number= 3
 XX FT 4170..5068
 XX FT /tag= g
 XX FT /number= 3
 XX FT 5069..5197
 XX FT /tag= h
 XX FT /number= 4

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 317.549 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51

Sequence: 1 catggcgccacggcgccacgc.....ggactctggcccccagtgtga 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl1:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	51	100.0	194	6	A57715
2	51	100.0	194	6	AR175909
3	51	100.0	13011	6	E14395
4	51	100.0	13011	6	AX827302
5	51	100.0	13011	10	RNLPKG
6	43.2	84.7	2717	10	RATPKRL1
7	43.2	84.7	231241	2	AC097039
8	36.6	71.8	415	9	HSPYKINPR
9	36.6	71.8	8409	9	HSU47654
10	36.6	71.8	13322	9	AY316591
11	36.6	71.8	13325	9	AL713999
12	33.6	65.9	192549	2	AC148047
13	32.2	63.1	95718	2	AC147942
14	30.2	59.2	136883	2	AC147944
15	30.2	59.2	153222	2	AC148054
16	30	58.8	136377	2	AC148066
17	30	58.8	174734	2	AC148068
18	28.8	56.5	178895	2	AC148063
19	25.8	50.6	140869	2	AC148210

c 20	25.8	50.6	166576	2	AC148209	AC148209	Macropus
c 21	25.2	49.4	158643	2	AC143451	AC143451	Macaca mu
c 22	25.2	49.4	174642	2	AC143448	AC143448	Macaca mu
c 23	25.2	49.4	239210	2	AC143094	AC143094	Macaca mu
c 24	25.2	49.4	349323	1	BX572096	BX572096	Prochiloro
c 25	25	49.0	149618	9	AP000556	AP000556	Homo sapi
c 26	25	49.0	150036	9	AP000557	AP000557	Homo sapi
c 27	25	49.0	166861	2	AC148587	AC148587	Homo sapi
c 28	25	49.0	167076	2	AC136511	AC136511	Homo sapi
c 29	25	49.0	216051	9	AC023490	AC023490	Homo sapi
c 30	25	49.0	212651	9	AC007957	AC007957	Homo sapi
c 31	25	48.2	173153	9	AC008119	AC008119	Homo sapi
c 32	24.6	48.2	179773	2	AC115041	AC115041	Papio ham
c 33	24.6	48.2	188462	2	AL592301	AL592301	Human DNA
c 34	24.6	48.2	248002	2	AC098668	AC098668	Human DNA
c 35	24.4	47.8	1255	9	HSU79264	HSU79264	Human clone
c 36	24.4	47.8	102386	4	AC091404	AC091404	Sus scrofa
c 37	24.4	47.8	151453	2	AC118988	AC118988	Sus scrofa
c 44	24.4	47.8	174559	2	AC027060	AC027060	Homo sapi
c 45	24.4	47.8	179714	9	AC092959	AC092959	Homo sapi

ALIGNMENTS

RESULT 1	A57715	Sequence 1 from Patent WO9632489.	194 bp	DNA	linear	PAT 03-MAR-1998
LOCUS	A57715					
DEFINITION	A57715	Sequence 1 from Patent WO9632489.				
ACCESSION	A57715					
VERSION	A57715.1	GI:3713539				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE	1	unclassified				
AUTHORS	Chen, R., Doiron, B. and Kahn, A.					
TITLE	GLUCOSE-INDUCIBLE RECOMBINANT VIRAL VECTOR					
JOURNAL	Patent: WO 9632489-A 1 17-OCT-1996;					
COMMENT	INST NAT SANTE RECH MED (FR)					
FEATURES	Other publication AU 5652396 961030					
source	Other publication FR 2732978 961018.					
location/Qualifiers						
1..194						
/organism="unclassified"						
/mol_type="unassigned DNA"						
/db_xref="taxon:32644"						
ORIGIN						
Query Match	100.0%;	Score 51;	DB 6;	Length 194;		
Best Local Similarity	100.0%;	Pred. No. 2;	7e-06;			
Matches	51;	Conservative	0;	Mismatches	0;	Indels
0;	Gaps	0;				
Qy	1	CATGGCGCCACGGGCACTCCCGTGTCTGACTGTGCCCCAGTGTGA 51				
Db	11	CATGGCGCCACGGGCACTCCCGTGTCTGACTGTGCCCCAGTGTGA 61				
RESULT 2	AR175909	Sequence 1 from patent US 6309878.	194 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR175909					
DEFINITION	AR175909	Sequence 1 from patent US 6309878.				
ACCESSION	AR175909					
VERSION	AR175909.1	GI:17917208				
KEYWORDS						
SOURCE						
ORGANISM						
Unknown.						

Unclassified.
REFERENCE 1 (bases 1 to 194)
AUTHORS Chen,R., Doiron,B. and Kahn,A.
TITLE Glucose-inducible recombinant viral vector
JOURNAL Patent: US 6309878-A 1 30-OCT-2001;
FEATURES Location/Qualifiers
source 1..194
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTG3CCCCCAGTGTGA 51
|||||
11 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTG6CCCCCAGTGTGA 61

Db
11 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTG6CCCCCAGTGTGA 61

RESULT 3
E14395 13011 bp DNA linear PAT 28-JUL-1999
LOCUS gDNA encoding pyruvate kinase.
DEFINITION E14395
ACCESSION E14395.1 GI:5709078
VERSION JP 1997313059-A/14.
KEYWORDS Ratius norvegicus (Norway rat)
SOURCE Ratius norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 13011)
REFERENCE 1
AUTHORS Muraase,M., Muraase,J., Iwabuchi,M., Hayakawa,T. and Imamura,J.
TITLE INCREASE IN STORED LIPID CONTENT OF PLANT SEED
JOURNAL Patent: JP 1997313059-A 14 09-DEC-1997;
MITSUBISHI CORP, MITSUBISHI CHEM CORP
OS Ratius norvegicus (rat)
PN JP 1997313059-A/14
PD 09-DEC-1997
PE 31-JAN-1997 JP 1997018966
PR 01-FEB-1996 JP 96P 16590
PI MPBASE MAKOTO, MURASE JUNKO, IWABUCHI MARI, HAYAKAWA TAKAHIKO,
PI IMAMURA JUN
PC A01H5/00, C07H21/04, C12N5/10, C12N9/12, C12N15/09, (C12N5/10, PC
C12N1:91),
PC (C12N9/12, C12R1:91), (C12N15/09, C12R1:91);
CC strandness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..13011
FT 3194.. .3218
/organism='Rattus norvegicus' FT exon
/number=1
FT intron 3219.. .3765
/number=1
FT exon 3766.. .3948
/number=2
FT intron 3949.. .5916
/number=2
FT exon 5917.. .6008
/number=3
FT intron 6009.. .6151
/number=3
FT exon 6152.. .6283
/number=4
FT intron 6284.. .6417
/number=4
FT exon 6418.. .6604
/number=5

FT intron 6605.. .6920
/number=5
FT exon 6921.. .7191
/number=6
FT intron 7192.. .7301
/number=6
FT exon 7302.. .7452
/number=7
FT intron 7453.. .7664
/number=7
FT exon 7665.. .7817
/number=8
FT intron 7818.. .7910
/number=8
FT exon 7911.. .8077
/number=9
FT intron 8078.. .9297
/number=9
FT exon 9298.. .9479
/number=10
FT intron 9480.. .10162
/number=10
FT exon 10163.. .10269
/number=11
FT 5'UTR 1.. .3193
FT 3'UTR 10270.. .13011
FT CDS
6152.. .6283 join(3212.. .3218,3766.. .3948,5917.. .6008, FT
6418.. .6604,6921.. .7191,7302.. .7452, FT
7665.. .7817,7911.. .8077,
9298.. .9479,10163.. .11594)
/product='Pyruvate kinase'.
FT
/location/Qualifiers
1..13011
/organism='Rattus norvegicus'
/mol_type='genomic DNA'
/db_xref='taxon:10116'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 13011;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTG6CCCCCAGTGTGA 51
|||||
Db 3021 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTG6CCCCCAGTGTGA 3071

RESULT 4
AX827302 13011 bp DNA linear PAT 12-DEC-2003
DEFINITION AX827302
ACCESSION AX827302
VERSION AX827302.1 GI:39837391
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE 1
AUTHORS Boess,F., Suter-Dick,L. and Wolf,D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 36 17-SEP-2003;
F. HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers
source 1..13011
/organism='Rattus norvegicus'
/mol_type='unassigned DNA'
/db_xref='taxon:10116'

ORIGIN

Query Match 100.0%; Score 51; DB 10; Length 13011;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 51
 |||||
 |||||
 DB 3021 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 3071

RESULT 6
 RATPKRL1 2717 bp DNA linear ROD 27-APR-1993
 LOCUS Rat pyruvate kinase gene, exons 1 and 2.
 DEFINITION M17088 J03455
 ACCESSION M17088.1 GI:206208
 VERSION
 KEYWORDS pyruvate kinase.
 SEGMENT 1 of 4
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2717)
 AUTHORS Noguchi,T., Yamada,K., Inoue,H., Matsuda,T. and Tanaka,T.
 TITLE The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters
 JOURNAL J. Biol. Chem. 262 (29), 14366-14371 (1987)
 MEDLINE 88007696
 PUBMED 3654663

COMMENT Original source text: Rat DNA (library of J. Bonner), clone lambda-LPK30, and reticulocyte, cDNA to mRNA, clones pRPK(1,3).
 Draft entry and printed copy of sequence for [1] kindly provided by T.Noguchi, 09-SEP-1987.

FEATURES
 source Location/Qualifiers
 1..2717

/organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 prim_transcript 925..2717
 /note="PK mRNA and introns (alt.)"
 prim_transcript 930..2717
 /note="PK mRNA and introns (alt.)"
 prim_transcript 934..2717
 /note="PK mRNA and introns (alt.)"
 exon <980..1079
 /note="R-pyruvate kinase"
 intron 1080..2007
 /note="PK intron A"
 exon <1446..1452
 /note="L-pyruvate kinase"
 intron 1453..2007
 /note="PK intron A (alt.)"
 exon 2008..2190
 /number=2
 intron 2191..22717
 /note="PK intron B"
 ORIGIN 342 bp upstream of PstI site.

Query Match 84.7%; Score 43.2; DB 10; Length 2717;
 Best Local Similarity 93.8%; Pred. No. 0.00059;
 Matches 45; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 51
 |||||
 |||||
 DB 1258 GGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 1305

RESULT 7
 LOCUS AC097039 231241 bp DNA linear HTG 10-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-61E1, *** SEQUENCING IN PROGRESS ***.

ACCESSION AC097039
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 231241)
 AUTHORS Muzny,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceaar,H., Centre,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Duhdin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,J., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howell,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Kapachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Krief,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenzhuwa,L., Louised,H., Lorado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwokenemeh,O., Okwum,G., Olarinuwaogun,A., Pal,S., Parks,K., Patelrnak,S., Paul,H., Perez,A., Perez,L., Plannoch,C., Plopper,F., Polndexter,A., Popovic,D., Primus,E., Pu,L., L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Syatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,Y., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weise,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinsteu,G. and Gibbs,R.A.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 231241)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (06-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 231241)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On May 10, 2003 this sequence version replaced gi:24819342.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are oriented and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHYP

Center clone name: CH230-61E1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 210827 bases at least Q40

Consensus quality: 214055 bases at least Q30

Estimated insert size: 223098; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 231241: contig of 231241 bp in length.

----- Location/Qualifiers

1..231241

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-61E1"

1

/note="clone boundary"

clone_end:sp6

site:ECORI

end_sequence:BH282405"

755..1528

/note="clone boundary"

clone_end:T7

site:ECORI

end_sequence:BH282404"

66116..68518

/note="wgs_end_extension"

clone_end:T7"

68569..69982

/note="wgs_end_extension"

clone_end:T7"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

RESULT 8

HSPYKINPR

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (03-DEC-1992) de Medicis E., Universite de Sherbrooke, de

Biochimie, 12e avenue Nord, Fleurimont, Quebec, Canada, J1H 5N4

Location/Qualifiers

1..415

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/cell_type="Leukocyte"

1..412

/note="liver pyruvate kinase promoter region"

/citation=[1]

promoter

Query Match

Best Local Similarity

Matches

39; Conservative

0; Mismatches

4; Indels

0; Gaps

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

0;

RESULT 9

HSU47654

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (30-JAN-1996) Peter Nurnberg, Institute for Medical

Genetics, Charite Medical School of the Humboldt-University Berlin,

Schumannstr. 20/21, Berlin 10098, Germany

Location/Qualifiers

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

```

mRNA
      join(<1, .139,1145, .1327,2508, .2599,2696, .2827,2961, .3147,
3510, .3780,3877, .4027,4672, .4824,4919, .5085,6325, .6506,
7572, .7678)
/product="pyruvate kinase PK-R isoenzyme"
join(<1, .139,1145, .1327,2508, .2599,2696, .2827,2961, .3147,
3510, .3780,3877, .4027,4672, .4824,4919, .5085,6325, .6506,
7572, .7678)
/codon_start=1
/product="pyruvate kinase PK-R isoenzyme"
/protein_id="AA092535.1"
/db_xref="GI:1230589"
/translation="HSMVPOQAHTESSMIOENISLQIQRSWKSGORDLAKSILIIIGA
PGGAGYLRRASVAOLTELGTAFPOOOLPAMADTFLEHLCIDIDSEPARSTSIATIGPARSV
IATIGPARSVBLKEMTKAGMNTARLNFSGHSEYHSAANREANESAPSGPLS
YRPVALIDTKPEIRTGILQSGPSESVLVGSOVLVTVDPARTGNANTVWDYDIP
NIVRVVPGGRITIDGLISLVQKIGPGLVTVQENGVLSRKGVLPGAOVDPG
LSEQVDRDLRFVEHGVDFVFAFPAKASDVAVRALGPEGHGKIISKIENHGVK
RDEILVSDGIMVARGDLEIIPAKVFLAQKMGIRCNAGKPVCAQOMLESMIT
KRPRTAETSDVANNVLDGACIMLGETAKNPFPEAVQMOMAIAREAAVYHROL
FELRRAPLSDPTEVTAIGAVEAFAKCAAIIVLTGGRSAQLLSRYRRAVIA
VTRSAQAAROVHLCRGVPLLYRPEPALMADVDRRVQGIESKLRGFLRVGLVI
VTGWRPGSGYTINRVLSTIS"
join(<478, .553,1145, .1327,2508, .2599,2696, .2827,
2961, .3147,3510, .3780,3877, .4027,4672, .4824,4919, .5085,
6325, .6506,7572, .7678)
/product="pyruvate kinase PK-L isoenzyme"
join(478, .553,1145, .1327,2508, .2599,2696, .2827,2961, .3147,
3510, .3780,3877, .4027,4672, .4824,4919, .5085,6325, .6506,
7572, .7678)
/codon_start=1
/product="pyruvate kinase PK-L isoenzyme"
/protein_id="AA092536.1"
/db_xref="GI:1230590"
/translation="WELTTOQCGADPQRRPREVSGMEGPAQYLRASVAOLTELQ
TAFFOQOLPAMADTFLEHLCIDIDSEPARSTSIATIGPARSVBLKEMTKA
GMNIRLNFSGHSEYHSAANREANESAPSGPLSYPVALIDTKPEIRTGILQ
GGPSELELVKGSQVLVTVDPARTGNANTVWDYDIPNIVRVVPGGRITIDGLISL
VYQKIGPGLVTVQENGVLSRKGVLPGAOVDPGLSEQVDRDLRFVEHGVDFV
ASFVKASDVAVRALGPEGHGKIISKIENHGVKRPDEILVSDGIMVARGDGI
EIPAKVFLAQKMGIRCNAGKPVCAQOMLESMITKRPRTAETSDVANNVLDGAC
IMLGETAKNPFPEAVQMOMAIAREAAVYHROLFELRRAPLSDPTEVTAIGAVEA
FKAACAAIIVLTGGRSAQLLSRYRRAVIAVTRSAQAAROVHLCRGVPLLY
RPEPALMADVDRRVQGIESKLRGFLRVGLVIIVTGWPRPGSGYTINRVLSTIS"

ORIGIN
Query Match      71.8%; Score 36.6; DB 9; Length 8409;
Best Local Similarity 90.7%; Pred. No. 0,068;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      9 CACGGGACACCTCCGCTGACTGACCTGCCCCCATGTA 51
      |||||||
Db      356 CACGGGACACCTCCGCTGACTGACTGCCCCCATGTA 398

RESULT 10
LOCUS      AY316591      13322 bp      DNA      linear      PRI 21-JUN-2003
DEFINITION      Homo sapiens pyruvate kinase, liver and RBC (PKLR) gene, complete cds.
ACCESSION      AY316591
VERSION      AY316591.1      GI:32140478
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 13322)
            Rieder,M.J., Livingston,R.J., Daniels,M.R., Montoya,M.A.,
            Chung,M.-W., Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L.,
            Robertson,F.D., Schackwitz,W.S., Sherwood,J.K., Wiltrak,L.A. and
            Nickerson,D.A.
```

```

TITLE      Direct Submission
JOURNAL
COMMENT      Submitted (06-JUN-2003) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
To cite this work please use: NIEHS-SNPs, Environmental Genome
Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
(URL: http://esg.gs.washington.edu).
FEATURES
source
1..13322
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..187
/notice="Region not scanned for variation"
misc_feature
9..157
/rpt_family="ERV"
/rpt_type=dispersed
variation
402
/frequency="0.01"
/replacement="c"
<525..11691
/gene="PKLR"
join(<525, .649,1665, .1847,6189, .6280,6377, .6508,
6643, .6829,7193, .7463,7560, .7710,8355, .8507,8602, .8768,
10008, .10189,11267, .11691)
/gene="PKLR"
/product="pyruvate kinase, liver and RBC"
join(550, .649,1665, .1847,6189, .6280,6377, .6508,6643,
7193, .7463,7560, .7710,8355, .8507,8602, .8768,10008, .10189,
11267, .11373)
/gene="PKLR"
/codon_start=1
/product="pyruvate kinase, liver and RBC"
/protein_id="AAP6927.1"
/db_xref="GI:32140479"
/translation="WSTIOENISLQIQRSWKSGORDLAKSILIIIGAPGAGYLRRASV
AOLQOELGTAFPOOOLPAMADTFLEHLCIDIDSEPARSTSIATIGPARSV
BLKEMTKAGMNTARLNFSGHSEYHSAANREANESAPSGPLSYPVALIDTKGP
EIRGILQSGPSESVLVGSOVLVTVDPARTGNANTVWDYDIPNIVRVVPGGRIT
IDGLISLVQKIGPGLVTVQENGVLSRKGVLPGAOVDPGLSEQVDRDLRFV
EHGVDIYASVVRKASDVAVRALGPEGHGKIISKIENHGVKRPDEILVSDGIM
VARQDLGLEIPAEVFLAQKMGIRCNAGKPVCAQOMLESMITKRPRTAETSDVA
NAVLDGACIMLGETAKNPFPEAVQMOMAIAREAAVYHROLFELRRAPLSDR
PTEVTAIGAVEAFAKCAAIIVLTGGRSAQLLSRYRRAVIAVTRSAQAAROVHL
CRGVPLLYRPEPALMADVDRRVQGIESKLRGFLRVGLVIIVTGWPRPGSGYT
INRVLSTIS"
repeat_region
1346..1567
/rpt_family="MIR"
/rpt_type=dispersed
variation
1572
/gene="PKLR"
/frequency="0.01"
/replacement="c"
1745
/gene="PKLR"
/frequency="0.01"
/replacement="c"
1905
/gene="PKLR"
/frequency="0.02"
/replacement="a"
1951..2139
/rpt_family="L2"
/rpt_type=dispersed
variation
1956
/gene="PKLR"
/frequency="0.03"
/replacement="c"
1960
/gene="PKLR"
/frequency="0.42"
/replacement="c"
2081..4840
/gene="PKLR"
```



```

repeat_region /note="Region not scanned for variation"
                2140..2451
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 2457..2758
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 2759..2810
                /rpt_family="L2"
                /rpt_type=dispersed
repeat_region 3003..3280
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3302..3318
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3319..3611
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3612..3909
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3920..5300
                /rpt_family="L2"
                /rpt_type=dispersed
variation      5122
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      5162
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      5223
                /gene="PKLR"
                /frequency="0.02"
                /replace="t"
repeat_region 5301..5619
                /rpt_family="Alu"
                /rpt_type=dispersed
variation      5401
                /gene="PKLR"
                /frequency="0.37"
                /replace="a"
repeat_region 5620..5705
                /rpt_family="L2"
                /rpt_type=dispersed
variation      5698
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
misc_feature  5890..5995
                /gene="PKLR"
                /note="Region not scanned for variation"
variation      6075
                /gene="PKLR"
                /frequency="0.03"
                /replace="g"
variation      6290
                /gene="PKLR"
                /frequency="0.02"
                /replace="t"
variation      6544
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
variation      6559
                /gene="PKLR"
                /frequency="0.40"
                /replace="c"
variation      7113
                /gene="PKLR"
                /frequency="0.01"

```

```

variation      7566
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
variation      7736
                /gene="PKLR"
                /frequency="0.01"
                /replace="a"
repeat_region 7788..8094
                /rpt_family="Alu"
                /rpt_type=dispersed
variation      7864
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      7898
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      7943
                /gene="PKLR"
                /frequency="0.01"
                /replace="g"
variation      8306
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      8326
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      8352
                /gene="PKLR"
                /frequency="0.01"
                /replace="t"
variation      8550
                /gene="PKLR"
                /frequency="0.01"

```

Query Match 71.8%; Score 36.6; DB 9; Length 13322;
 Best Local Similarity 90.7%; Pred. No. 0.063; 4; Indels 0; Gaps 0;
 Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Qy      9 CACGGGCACTCCCGTGTCTCTGACTCTGCCCCAGTGA 51
Db      867 CACGGGCACTCCCGTGTCTCTGACTCTGCCCCAGTGA 909

```

```

RESULT 11
LOCUS      AL713999/C      133525 bp      DNA      linear      PRI 13-MAY-2003
DEFINITION Human DNA sequence from clone Rpl1-263K19 on chromosome 1, complete
sequence.
ACCESSION  AL713999
VERSION    AL713999.28  GI:30722340
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 133525)
AUTHORS  Glithero,R.
TITLE    Direct Submission
JOURNAL  Submitted (13-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 14, 2003 this sequence version replaced gi:30523877.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl> RP11-263K19 is from the library RPCT-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.
 Location/Qualifiers
 1..133525
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-263K19"
 /clone_1ib="RPCT-11.1"

ORIGIN

Query Match 71.8%; Score 36.6; DB 9; Length 133525;
 Best Local Similarity 90.7%; Pred. No. 0.041;
 Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGGACCTCCCGTTCTGACCTCGGCGCCCACTGTA 51
 |||||
 Db 129358 CACGGGGACCTCCCGTTCTGACCTCGGCGCCCACTGCA 129316

RESULT 12
 AC148047/c
 LOCUS
 DEFINITION Otollemur garnettii clone CH256-360P11, WORKING DRAFT SEQUENCE, 7
 ORDERED PIECES.
 AC148047
 AC148047.2 GI:45544599
 HTG: HTGS PHASE2; HTGS DRAFT.
 SOURCE Otollemur garnettii (small-eared galago)
 ORGANISM Otollemur garnettii
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Strepsirrhini; Galagonidae; Otollemur.
 1 (bases 1 to 192549)
 Antonellis, A., Ayele, K., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B., Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P., Larson, S., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Marulles, E.H., Mastaglio, C., Maskeri, B., McQuell, J., Mullikin, J.C., Pagnigan, C., Portnoy, M.E., Prasad, A., Puri, O., Redix-Dugue, N., Schandler, K., Schneider, M.G., Shah, K., Sison, C., Statistop, S., Thomas, J.W., Thomas, P.J., Tsiouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 TITLE
 JOURNAL
 Unpublished

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 2 (bases 1 to 192549)
 Green, E.D.
 Direct Submission
 Submitted (29-JAN-2004) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 192549)
 Green, E.D.
 Direct Submission
 Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Mar 18, 2004 this sequence version replaced gi:41386816.

 Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov

 Project Information
 Center project name: fxd
 Center clone name: 360P11

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 190022 bases at least Q40
 Consensus quality: 191157 bases at least Q30
 Consensus quality: 191709 bases at least Q20
 Insert size: 149000; agarose-1p
 Insert size: 191949; sum-of-contigs
 Quality coverage: 11.55x in Q20 bases; agarose-1p
 Quality coverage: 8.97x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 38219: contig of 38219 bp in length
 * 38220 38319: gap of unknown length
 * 38320 77219: contig of 38900 bp in length
 * 77220 77319: gap of unknown length
 * 77320 114115: contig of 36796 bp in length
 * 114116 114215: gap of unknown length
 * 114216 126274: contig of 12059 bp in length
 * 126275 126374: gap of unknown length
 * 126375 152697: contig of 26323 bp in length
 * 152698 152977: gap of unknown length
 * 152978 166906: contig of 14109 bp in length
 * 166907 167006: gap of unknown length
 * 167007 192549: contig of 25543 bp in length.

FEATURES

source
 1..192549
 /organism="Otollemur garnettii"
 /mol_type="genomic DNA"
 /db_xref="taxon:30611"
 /clone="CH256-360P11"
 /clone_1ib="CH256"

```

misc_feature      /note="BAC resource: http://bacpac.chori.org/"
                  1..38219
                  /note="assembly_fragment
                  clone_end:SP6
                  vector_side:left"
                  1..30749
misc_feature      /note="Clone overlaps with Genbank Accession Number
                  AC148125 clone CH256-404N19 (center project name fix)"
                  38320..77219
                  /note="assembly_fragment"
                  77320..114115
                  /note="assembly_fragment"
                  114216..126274
                  /note="assembly_fragment"
                  126375..152697
                  /note="assembly_fragment"
                  152798..166906
                  /note="assembly_fragment"
                  167007..192549
                  /note="assembly_fragment
                  clone_end:T7
                  vector_side:right"

ORIGIN
Query Match      65.9%; Score 33.6; DB 2; Length 192549;
Best Local Similarity 81.2%; Pred.No.0.36;
Matches 39; Conservative 0; Mismatches 9; Indels 0; Gaps 0

Qy               4 GGGCGCAGCGGACACTCCCGTGTTCCTCGAGACTGTGCCCCAGATGA 51
                  |||
Db               94043 GGAGCGACGGGGCATTCCTCGTGTCTCCCTCGAGACTTGGCCCCCTGCTTA 93996

RESULT 13
AC147942/c
LOCUS              AC147942                95718 bp    DNA          linear   HTG 18-MAR-2004
DEFINITION         Carollia peripicillata clone 47203, WORKING DRAFT SEQUENCE, 5
ordered pieces.
ACCESSION          AC147942
VERSION            AC147942.2 GI:45544627
KEYWORDS           HTG, HTGS PHASE2, HTGS DRAFT.
SOURCE             Carollia peripicillata (Seba's short-tailed bat)
ORGANISM           Carollia peripicillata
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
                  Carolliinae; Carollia.
REFERENCE          1 (bases 1 to 95718)
AUTHORS            Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
                  Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
                  Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
                  Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
                  Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,
                  Larson,S., Lee-Jin,S.-O., Leggett,R., Maduro,Q.L., Maduro,V.B.,
                  Margulies,E.H., Masiello,C., Mawkeri,B., McDowell,J.,
                  Mullikin,J.C., Pagnutigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
                  Reddi-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
                  Stentropop,S., Thomas,J.W., Thomas,P.J., Telpouri,V., Vogt,J.L.,
                  Wehrhoy,K.D., Young,A. and Green,E.D.
TITLE              NISC Comparative Sequencing Initiative
JOURNAL            Unpublished
REFERENCE          2 (bases 1 to 95718)
AUTHORS            Green,E.D.
TITLE              Direct Submission
JOURNAL            Submitted (22-JUN-2004) NIH Intramural Sequencing Center, 8717
                  Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE          3 (bases 1 to 95718)
AUTHORS            Green,E.D.
TITLE              Direct Submission
JOURNAL            Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717
                  Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT            On Mar 18, 2004 this sequence version replaced gi:41058168.
                  ----- Genome Center
                  Center: NIH Intramural Sequencing Center

```

```

Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoomhgri.nih.gov
----- Project Information -----
Center project name: eju
Center clone name: 472003

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones; overlaps with neighboring
clones; alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 94695 bases at least Q40
Consensus quality: 95070 bases at least Q30
Consensus quality: 95243 bases at least Q20
Insert size: 91000; agarose-fp
Insert size: 95318; sum-of-contigs
Quality coverage: 11.68x in Q20 bases; agarose-fp
Quality coverage: 11.15x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1      18808: contig of 18808 bp in length
*
* 18809      18908: gap of unknown length
*
* 18909      22119: contig of 3211 bp in length
*
* 22120      22219: gap of unknown length
*
* 22220      31750: contig of 9531 bp in length
*
* 31751      31850: gap of unknown length
*
* 31851      53063: contig of 22123 bp in length
*
* 53064      53163: gap of unknown length
*
*          95718: contig of 42555 bp in length.
*
* Location/Qualifiers
*   1..95718
*     /organism="Carollia peroplicillata"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:40233"
*     /clone_1fb="47203"
*     /clone_1lb="The Green Lab"
*     /note="greenohgri.nih.gov"
*
* misc_feature
*   1..18808
*     /note="assembly_fragment
* clone_end:S6
* vector_side:left"
*
* misc_feature
*   1..4885
*     /note="clone overlaps with GenBank Accession Number
AC148052 clone 593F8 (center project name gda)"
*
* misc_feature
*   18909..22119
*     /note="assembly_fragment"
*
* misc_feature
*   22220..31750
*     /note="assembly_fragment"
*
* misc_feature
*   31851..53063
*     /note="assembly_fragment"
*
* misc_feature
*   53164..95718
*     /note="assembly_fragment
* clone_end:T7

```


misc_feature 139917..148451 /note="assembly_fragment"
misc_feature 148552..153222 /note="assembly_fragment"

clone_end:SP6
vector_side:right"

ORIGIN

Query Match 59.2%; Score 30.2; DB 2; Length 153222;

Best Local Similarity 81.4%; Pred.No. 4.9;

Matches 35; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ATGGGGGACACGGGGACATCCCGTGGTTCCTGGACTCTGGCCCC 44
Db 68059 ATGGGGCCATGGGACAGTCCCATGTCTCCTGGACTCTGGCCCC 68017

Search completed: September 1, 2005, 03:26:37
Job time : 327.549 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 387.736 Seconds
(without alignments)
5006.706 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51
Sequence: 1 catggggcgcacggggcactc.....ggacctggccccctgtcta 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsb1:*
9: gb_gsb2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	84.3	659	8	AZ840793 2M0138D02
2	28.8	56.5	406	4	BM286990 527229 MA
3	27.6	54.1	827	4	BF105536 601822455
4	26.4	51.8	907	2	BE867308 601442381
5	26	51.0	765	1	AL528750 AL528750
6	25.8	50.6	379	9	CE316355 tigr-gsb-
7	25.8	50.6	672	9	CE450351 tigr-gsb-
8	25.8	50.6	739	7	CN360924 170004706
9	25.6	50.2	225	6	CD630056 560279323
10	25.6	50.2	229	6	CD630055 560279323
11	25.6	50.2	1021	5	BO056161 AGENCOURT
12	25.2	49.4	472	2	AM922227 DGI_17_E0
13	25	49.0	259	8	CC453282 ZMMBC034
14	25	49.0	334	4	BI006710 RCS-RT005
15	25	49.0	403	9	CE744237 tigr-gsb-
16	25	49.0	481	2	AM577607 PM2-BT055
17	25	49.0	966	5	BU186634 AGENCOURT
18	25	49.0	969	4	BG482095 602527983
19	24.6	48.2	865	8	BZ404685 OGAAM77TC
20	24.6	48.2	1193	5	BQ438699 AGENCOURT
21	24.4	47.8	416	9	CE608928 tigr-gsb-
22	24.4	47.8	416	9	CE565764 tigr-gsb-
23	24.4	47.8	460	2	BF189053 234314 MA
24	24.4	47.8	844	1	AU079412 AU079412

C	25	24.4	47.8	884	2	BF101581 601753396
	26	24.4	47.8	899	9	CC972844 ZUACY29TH
	27	24.4	47.8	913	5	BQ940337 AGENCOURT
C	28	24.2	47.5	265	9	CE418242 tigr-gsb-
	29	24.2	47.5	267	9	CG481325 OST13706
C	30	24.2	47.5	377	6	CB177707 tn94f04.x
	31	24.2	47.5	413	7	CO739485 SLB06A23
	32	24.2	47.5	442	9	CE200317 tigr-gsb-
C	33	24.2	47.5	469	9	CE699841 tigr-gsb-
	34	24.2	47.5	518	7	CN669363 A0878H10-
	35	24.2	47.5	521	5	BM944811 UI-M-EHD
	36	24.2	47.5	523	9	CE338459 tigr-gsb-
	37	24.2	47.5	528	7	CN670899 A0898H12-
	38	24.2	47.5	539	5	BM943713 UI-M-EHD
	39	24.2	47.5	541	7	CE627560 1005f01.y
	40	24.2	47.5	552	9	CE668350 tigr-gsb-
	41	24.2	47.5	569	5	BM944273 UI-M-EHD
	42	24.2	47.5	572	9	CE075799 tigr-gsb-
	43	24.2	47.5	594	7	CK619547 mk25e02.y
	44	24.2	47.5	595	9	CE123626 tigr-gsb-
	45	24.2	47.5	610	9	CE256374 tigr-gsb-

ALIGNMENTS

RESULT 1
AZ840793 659 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0138D02R Mouse 10kb plasmid U08C1M library mus musculus genomic
LOCUS AZ840793
ACCESSION AZ840793
VERSION AZ840793.1 GI:13010701
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Dun, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
1 (bases 1 to 659)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0138 row: D column: 02
Seq primer: CACACAGCAACAGCTATGACC
Class: plasmid ends
High quality sequence atp: 659.
location/Qualifiers
1..659
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U08C2M0138D02"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid U08C1M library"
/note="Vector: pMD29v; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 84.3%; Score 43; DB 8; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.00051;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTCTCTGACTGTGCCCCCACTGTA 51
Db 434 CACGGGCACTCCCGTCTCTGACTGTGCCCCCACTGTA 476

RESULT 2 406 bp mRNA linear EST 28-DEC-2001
LOCUS BM286990
DEFINITION 527829 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM286990
VERSION BM286990.1 GI:17996016
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM

REFERENCE 1 (bases 1 to 406)
AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,

Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkung,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Petrea,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

TITLE
JOURNAL
MEDLINE
PubMed
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCACGACG
Plate: 135 row: G column: 20
Seq primer: ATTTAGTGACACTTAG.
Location/Qualifiers

FEATURES
source
1. 406
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3BOV"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."

ORIGIN

Query Match 56.5%; Score 28.8; DB 4; Length 406;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GGGGCGACGGGCACTCCCGTCTCTGACTGTGCCCCCACTGTA 51
Db 124 GGAGCGAAGGCGCACTCCCATGACTGACTGTGCCCCCTGTGCA 171

RESULT 3 827 bp mRNA linear EST 19-OCT-2000
LOCUS BF105536
DEFINITION 601822455R1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4045620 3',
mRNA sequence.
ACCESSION BF105536
VERSION BF105536.1 GI:10888062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE 1 (bases 1 to 827)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov

Plate: L1CM871 row: m column: 13
High quality sequence start: 28
High quality sequence stop: 34.
Location/Qualifiers
1. 827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4045620"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgcc); Site 2: SfiI (ggccatratggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

FEATURES
source

1. 827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4045620"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgcc); Site 2: SfiI (ggccatratggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 54.1%; Score 27.6; DB 2; Length 827;
Best Local Similarity 78.6%; Pred. No. 78;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 GGGGCGACGGGCACTCCCGTCTCTGACTGTGCCCCC 45
Db 404 GGGGCGAGGCGACACTCCCGGCGTTCGTGACGCTCGCCCC 363

RESULT 4 907 bp mRNA linear EST 20-OCT-2000
LOCUS BE867308


```

DEFINITION 601442381F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846429 5',
            mRNA sequence.
ACCESSION  BB667308
VERSION    BB667308.1 GI:10316084
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 907)
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cgabs@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM9559 row: a column: 22
            High quality sequence stop: 563.
FEATURES    Location/Qualifiers
            source          1..907
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:3846429"
                        /tissue_type="adenocarcinoma"
                        /lab_host="DH10B (phage-resistant)"
                        /note="Organ: Colon; Vector: PCMV-SPORT6; Site:1: NotI;
                        Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                        Average insert size 1.8 kb. Library constructed by Life
                        Technologies."
ORIGIN
Query Match          51.8%; Score 26.4; DB 2; Length 907;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 4 GGGCGACGGGACATCCCGTGTCTCTGACTCTGCCCCCAG 47
    |||||
DB 777 GGCTGCACGGGGCCCTGGCGGAGCATGTGTCTGCGACGAG 820

RESULT 5
AL528750 765 bp mRNA linear EST 24-MAR-2004
LOCUS    AL528750 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION
ACCESSION  AL528750
VERSION    AL528750
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 765)
REFERENCE  Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            On Feb 13, 2001 this sequence version replaced gi:31066600.
            Contact: Genoscope
            Genoscope - Centre National de Sequenage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: sequen@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a

```

division of Invitrogen. This sequence belongs to sequence cluster 6514.r and it belongs to a clone representative of this cluster. For more information about this cluster and the virtual cDNA, see <http://www.genoscope.cns.fr/cdna?e=CS0D0001CG04NP1&c=6514.r>.

```

FEATURES    Location/Qualifiers
            source          1..765
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0D0001YN07"
                        /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
                        /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
                        /note="1st strand cDNA was primed with a NotI-oligo(dT)
                        primer. Five prime end enriched, double-strand cDNA was
                        digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match          51.0%; Score 26; DB 1; Length 765;
Best Local Similarity 76.2%; Pred. No. 2.7e+02;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CATGGCGACGGGACATCCCGTGTCTCTGACTCTGGCC 42
    |||||
DB 445 CCTGGGACACTGTGCACTTCCCGAATCCTGGACTTGCC 486

RESULT 6
CE316355 379 bp DNA linear GSS 26-SEP-2003
LOCUS    CE316355/c
DEFINITION  tigr-gss-dog-17000361087224 dog library Canis familiaris genomic.
            genomic survey sequence.
ACCESSION  CE316355
VERSION    CE316355.1 GI:36119120
KEYWORDS   GSS.
SOURCE     Canis familiaris (dog)
ORGANISM   Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
            1 (bases 1 to 379)
REFERENCE  Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Ruesch,D.B., Delcher,A.L., Pop,M., Wang,W., Frazer,C.M. and
            Venter,J.C.
            The dog genome: survey sequencing and comparative analysis
            Science 301 (5641), 1898-1903 (2003)
            22875432
            MEDLINE 14512627
            PUBMED
            COMMENT  Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES    Location/Qualifiers
            source          1..379
                        /organism="Canis familiaris"
                        /mol_type="genomic DNA"
                        /strain="Standard Poodle"
                        /db_xref="taxon:9615"
                        /clone_lib="Dog Library"
                        /note="Site 1: BclXI; Libraries were prepared from
                        peripheral blood"
ORIGIN
Query Match          50.6%; Score 25.8; DB 9; Length 379;
Best Local Similarity 81.1%; Pred. No. 3e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGGCGACGGGACATCCCGTGTCTCTGACTC 37
    |||||

```

Db 346 CATGGGACAGTGGGCACTCCGCGCTCCCTGGGCTC 310

RESULT 7
LOCUS CE450351 672 bp DNA linear GSS 27-SEP-2003
DEFINITION tigr-gss-dog-17000319093723 Dog Library Canis familiaris genomic,
genomic survey sequence.

ACCESSION CE450351
VERSION CE450351.1 GI:36740297
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 672)
Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Frazer,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

TITLE JOURNAL
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Classes: shotgun.

FEATURES

source 1..672
Location/Qualifiers

/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 50.6%; Score 25.8; DB 9; Length 672;
Best Local Similarity 81.1%; Pred. No. 3.1e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CATGGCGGACAGTGGGCACTCCGCGCTCCCTGGGCTC 37
|||||
Db 553 CATGGGACAGTGGGCACTCCGCGCTCCCTGGGCTC 589

RESULT 8
LOCUS CN360924 739 bp mRNA linear EST 16-MAY-2004
DEFINITION 1700047066057 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN360924
VERSION CN360924.1 GI:47360858
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 739)
Brandenberger,R., Wei,H., Zhang,S., lei,S., Muraige,J., Fisk,G.J.,
Li,Y., Xu,C., Pang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J. and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6): 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA

TITLE JOURNAL
COMMENT

Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 739 Std Error: 0.00.
Location/Qualifiers
1..739

FEATURES

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 50.6%; Score 25.8; DB 7; Length 739;
Best Local Similarity 81.1%; Pred. No. 3.1e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 8 GCACGGGCACTCCGCGTCTCTGACTCTGGCCCC 44
|||||
Db 231 GCACAGGGCACTGCAGTGGGGCTGGCCTCTGGCCCC 195

RESULT 9
LOCUS CD630056 225 bp mRNA linear EST 12-JAN-2004
DEFINITION 5602793201 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD630056
VERSION CD630056.1 GI:40278322
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 225)
Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

ORIGIN

Query Match 50.2%; Score 25.6; DB 6; Length 225;
Best Local Similarity 77.5%; Pred. No. 3.4e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 GCACGGGCACTCCGCGTCTCTGACTCTGGCCCCAG 47
|||||
Db 85 GCACGTGACCAACCCGGGAGCTGTGACTCTGATCAG 124

FEATURES
source 1..225
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 50.2%; Score 25.6; DB 6; Length 225;
Best Local Similarity 77.5%; Pred. No. 3.4e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 GCACGGGCACTCCGCGTCTCTGACTCTGGCCCCAG 47
|||||
Db 85 GCACGTGACCAACCCGGGAGCTGTGACTCTGATCAG 124

RESULT 10
LOCUS CD630055 229 bp mRNA linear EST 12-JAN-2004
DEFINITION 5602793201 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD630055
VERSION CD630055.1 GI:40278321
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES
source
1..229
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1fb="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 50.2%; Score 25.6; DB 6; Length 229;
Best Local Similarity 77.5%; Pred. No. 3.4e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 GCACGGGGGACATCCCGTGGTCTCTGACTGTGCCCCAG 47
|||||
Db 140 GCACGTGACACCCGGGAGCCTGACTCTGCAATCCAG 101
|||||

RESULT 11
LOCUS BQ056161 1021 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT 6773628 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5808449
5', mRNA sequence.
ACCESSION BQ056161
VERSION BQ056161.1 GI:19815501
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb@r-mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2052 row: h column: 18
High quality sequence stop: 669.
Location/Qualifiers
1..1021
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1fb="IMAGE:5808449"
/issue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1fb="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7, Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

ORIGIN
Query Match 50.2%; Score 25.6; DB 5; Length 1021;
Best Local Similarity 77.5%; Pred. No. 3.7e+02;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 GGGCGACGGGACATCCCGTGGTCTCTGACTGTGCCCC 43
|||||
Db 988 GGGGCGAGGGGTACACCCGTGTTCTTGCGGCGTCCCC 949
|||||

RESULT 12
LOCUS AW922227/c 472 bp mRNA linear EST 19-JUL-2000
DEFINITION DG1_17_E04_b1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION AW922227
VERSION AW922227.1 GI:8088052
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Sorghum.
AUTHORS Cordonier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@pratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 420
POLYA=NO.
Location/Qualifiers
1..472
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_1fb="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN
Query Match 49.4%; Score 25.2; DB 2; Length 472;
Best Local Similarity 71.7%; Pred. No. 4.8e+02;
Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 4 GGGCGACGGGACATCCCGTGGTCTCTGACTGTGCCCCAGTG 49
|||||
Db 459 GTGACACGGGCGACATCGAGCAGCTGTGAACGCTGTGCGCGTG 414
|||||

RESULT 13
LOCUS CC453282/c 259 bp DNA linear GSS 22-MAY-2003
DEFINITION ZMMBBc0344B20r ZMMBBc Zea mays genomic clone ZMMBBc0344B20 3', genomic survey sequence.
ACCESSION CC453282

```

VERSION      CC453282.1  GI:31000845
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 259)
AUTHORS      Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
             Rouzard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE        Sequencing of the maize genome at PGR (2003b)
JOURNAL      Unpublished (2003)
COMMENT      Contact: Bharti,A.K.
             Dr.Joachim Messing's lab
             The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
             University
             190 Frelinghuysen Road, Piscataway, NJ 08854, USA
             Tel: 732 445 3801
             Fax: 732 445 5735
             Email: bharti@waksman.rutgers.edu
             Seq primer: SP6
             Class: BAC ends
             High quality sequence start: 83.
FEATURES     source
             1..259
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /culivar="B73"
             /db_xref="taxon:4577"
             /clone="ZMMBRC0344B20"
             /lab_host="E. coli DH10B"
             /clone_lib="ZMMBRC"
             /note="Vector: pTARBAC1.3; Site_1: BamHI; site_2: BamHI"

ORIGIN
Query Match      49.0%; Score 25; DB 8; Length 259;
Best Local Similarity 69.4%; Pred. No. 5.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY
3 TGGGGCGACGGGCACTCCGTCGTCGACCTGCGCCCGCAGTGA 51
133 TCGGCTACGGGAGAGCCCGAGTGTCTGACGTGCGCTCCCGCGTA 75

Db
1006710/c      334 bp      mRNA      linear      EST 13-JUN-2001
LOCUS          B1006710
DEFINITION    RCS-RT0053-110101-011-H09 RT0053 Homo sapiens CDNA, mRNA sequence.
ACCESSION     B1006710
VERSION       B1006710.1  GI:14410781
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 334)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
             Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
             Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
             Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
             O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
             Simpson,A.J.
TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
             sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
PUBMED       10737800
COMMENT      Contact: Simpson A.J.G.
             Laboratory of Cancer Genetics
             Ludwig Institute for Cancer Research
             Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
             Brazil
             Tel: +55-11-2704922

```

```

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
http://www.ludwig.org.br/scripts/gethtml2.pl?l=RCS&t2=RCS-RT0053-
110101-011-H09&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 334.
FEATURES     source
             1..334
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="RT0053"
             /note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;
             Site 2: SmaI; A mini-library was made by cloning products
             derived from ORESTS PCR (U.S. Letters Patent application
             No. 196,716 - Ludwig Institute for Cancer Research)
             profiles into the pUC 18 vector. Reverse transcription of
             tissue mRNA and cDNA amplification were performed under
             low stringency conditions."

ORIGIN
Query Match      49.0%; Score 25; DB 4; Length 334;
Best Local Similarity 75.6%; Pred. No. 5.5e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY
8 GCACGGGGCACTCCCGTCGTCCTGACCTGCGCCCGCAGT 48
133 GCACGGGGCTGACCGCTGCTCCCTGCGCTCGTTCACAGT 293

Db
CE744237/c      403 bp      DNA      linear      GSS 30-SEP-2003
LOCUS          CE744237
DEFINITION    tigr-gss-dog-17000369534847 Dog library Canis familiaris genomic,
             genomic survey sequence.
ACCESSION     CE744237
VERSION       CE744237.1  GI:37084584
KEYWORDS      GSS.
SOURCE        Canis familiaris (dog)
ORGANISM      Canis familiaris
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE     1 (bases 1 to 403)
AUTHORS      Kirkness,E.F., Batra,V., Halpern,A.L., Levy,S., Remington,K.,
             Ruesch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
             Venter,J.C.
TITLE        The dog genome: survey sequencing and comparative analysis
JOURNAL      Science 301 (5641), 1898-1903 (2003)
MEDLINE      22875432
PUBMED       14512627
COMMENT      Contact: Kirkness EF
             The Institute for Genomic Research
             Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
             Rockville, MD 20850, USA
             Tel: 301-838-0200
             Fax: 301-838-0208
             Email: ekirknes@tigr.org
             Class: shotgun.
FEATURES     source
             1..403
             /organism="Canis familiaris"
             /mol_type="genomic DNA"
             /strain="Standard Poodle"
             /db_xref="taxon:9615"
             /clone_lib="Dog Library"
             /note="Site 1: BstXI; Libraries were prepared from
             peripheral blood"

ORIGIN
Query Match      49.0%; Score 25; DB 9; Length 403;

```

Best Local Similarity 69.4%; Pred. No. 5.6e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CATGGGCGACGGGGCACTCCCGTGGTTCCTGAGCTCTGGCCCCCAGTG 49
Db 126 CTTGGGCTCAGGGGTGAATCCAGAGTTCCTGGAGTGAAGTCCACATTG 78

Search completed: September 1, 2005, 06:56:38
Job time : 399.902 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 12:52:03 ; Search time 54.0797 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51
Sequence: 1 catggcgccagcgagcagcactc.....gagactcgccccagctgta 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	51	ABX15371	Abx15371 Rat glucoc
2	51	100.0	194	AAT43002	Abx15373 Rat type
3	51	100.0	270	ABX15373	Abx15373 Rat insul
4	51	100.0	321	ABX15374	Abx15374 Rat insul
5	51	100.0	372	ABX15375	Abx15375 Rat insul
6	51	100.0	423	ABX15376	Abx15376 Rat insul
7	51	100.0	13011	2 AAT96631	Aat96631 CDNA enco
8	51	100.0	13011	10 AAT96631	Aat96631 CDNA enco
9	51	100.0	13011	12 AAT96631	Aat96631 CDNA enco
10	48	94.1	48	ABX15379	Abx15379 Rat liver
11	48	94.1	51	ABX15380	Abx15380 Rat liver
12	30	58.9	30	ABX15380	Abx15380 Rat liver
13	27	52.9	27	ABX15380	Abx15380 Rat liver
14	27	52.9	27	ABX15380	Abx15380 Rat liver
15	26	51.0	26	ABX15380	Abx15380 Rat liver
16	24.6	48.2	55001	12 ADG88848	Adg88848 Human Not
17	24.6	48.2	55001	12 ADH74825	Adh74825 Human Not
18	24.4	47.8	3138	6 AA62663	Aa62663 Human Zic
19	24.4	47.8	3138	6 AEN95218	Abn95218 Gene #171
20	24.4	47.8	3138	12 ADN04235	Adn04235 Antispori

ALIGNMENTS

21	24.4	47.8	3138	12	ADQ17359	Adq17359 Human bof
22	24.4	47.8	3138	13	ADR25048	Adr25048 Breast ca
23	24.4	47.8	3511	12	ADQ22099	Adq22099 Human bof
24	24.4	47.1	24	2	AAT43003	Aat43003 Rat type
25	24	47.1	209	12	ACH93032	Ach93032 Human gen
26	24	47.1	525	12	ACH79332	Ach79332 Human gen
27	24	47.1	712	4	AAK77580	Aak77580 Human imm
28	24	47.1	840	6	AAK77580	Aak77580 Human imm
29	24	47.1	1188	8	ABV73365	Abv73365 Human TGR
30	24	47.1	1191	6	ABZ42892	Abz42892 Human GPC
31	24	47.1	1191	12	ADO29943	Ado29943 Human GPC
32	24	47.1	1227	6	AAI68793	Aai68793 Human MAR
33	24	47.1	1311	9	ACC84331	Acc84331 Human hpl
34	24	47.1	1416	4	AAK77579	Aak77579 Human imm
35	24	47.1	1498	2	AAV38513	Aav38513 Human 7-c
36	24	47.1	1546	6	AAV98070	Aav98070 Human DNA
37	24	47.1	1546	8	ABZ42551	Abz42551 Human G P
38	24	47.1	1546	10	ABT17029	Abt17029 Human MP2
39	24	47.1	1579	4	AAI99548	Aai99548 Human exp
40	24	47.1	1595	4	AAH78712	Aah78712 Human EX3
41	24	47.1	1905	10	ADF70603	Adf70603 Orphan re
42	24	47.1	2046	4	AAZ28996	Aaz28996 Genomic s
43	24	47.1	2046	4	AAZ30237	Aaz30237 DNA encod
44	24	47.1	2046	4	AAZ35064	Aaz35064 DNA #14 e
45	24	47.1	2046	4	AAK77581	Aak77581 Human imm

RESULT 1
ID ABX15371 standard; DNA; 51 BP.
XX
XX ABX15371;
AC
XX
XX 17-APR-2003 (first entry)
DT
XX
XX Rat glucose response element (GIRE) DNA.
DE
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hypoglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS
XX
XX US2002107198-A1.
FN
XX
XX 08-AUG-2002.
PD
XX
XX 10-OCT-2001; 2001US-00972916.
PF
XX
XX 11-OCT-2000; 2000US-0239113P.
PR
XX
XX (THUL/) THUL P M.
PA
XX
XX Thule PM;
PI
XX
XX WPI; 2002-674190/72.
DR
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 7; Page 13; 37pp; English.
PS
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth

CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications (without severe hypoglycaemia), to
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat glucose response element of the invention
XX
SQ Sequence 51 BP; 6 A; 18 C; 17 G; 10 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CATGGCGCAGCGGGCACTCCGTTGTTCTGGACTCTGCCCCCAGTGTGA 51
Db 1 CATGGCGCAGCGGGCACTCCGTTGTTCTGGACTCTGCCCCCAGTGTGA 51
RESULT 2
AA143002
ID AA143002 standard; DNA; 194 BP.
XX
AC AA143002;
XX
DT 16-JUL-1997 (first entry)
XX
DE Rat type I pyruvate kinase promoter region.
XX
KW Glucose-inducible; rat; pyruvate kinase type I; L-PK; promoter;
KW hyperglycaemia; diabetes; gene therapy; viral vector; defective virus;
KW ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Promoter 16..39
FT /tag= a
FT /label= L4_element
FT /note= "able to interact with MLTF/USF factor"
FT promoter 40..58
FT /*tag= b
FT /label= L3_element
FT /note= "able to interact with HNF 4 factor"
XX
PN WO9632489-A1.
XX
PD 17-OCT-1996.
XX
PF 12-APR-1996; 96WO-FR000560.
XX
PR 14-APR-1995; 95FR-00004558.
XX
PA (INRM) INST NAT SANTE & RECH MEDICALE.
XX
PI Chen R, Doiron B, Kahn A;
XX
DR WPI; 1996-477139/47.
XX
PT Recombinant defective virus contg. heterologous gene under control of
PT glucose inducible promoter - and infected mammalian cells, esp. for gene
PT therapy of diabetes.
XX
PS Claim 4; Page 21; 41pp; French.

XX
CC A new defective recombinant virus contains at least one heterologous gene
CC under control of an expression signal inducible by glucose or its
CC analogues. Preferably, the expression signal is derived from the 183 bp
CC region located 5' of the type I pyruvate kinase (L-PK) gene, especially
CC having the present sequence. Expression of the recombinant virus can be
CC controlled by glucose levels, but insulin production can be stopped
CC immediately by administering glucagon. This avoids the risk of
CC hypoglycaemia caused by excessive insulin production. Mammalian cells
CC infected by the virus can be implanted into liver, spleen, pancreas or
CC intestine, to provide insulin secretion at predetermined sites. The virus
CC is useful in gene therapy to treat and/or prevent diseases associated
CC with hyperglycaemia, particularly diabetes, but more generally can be
CC used to provide controlled expression of a wide variety of proteins
XX
SQ Sequence 194 BP; 50 A; 59 C; 51 G; 34 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 2; Length 194;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CATGGCGCAGCGGGCACTCCGTTGTTCTGGACTCTGCCCCCAGTGTGA 51
Db 1 CATGGCGCAGCGGGCACTCCGTTGTTCTGGACTCTGCCCCCAGTGTGA 61
RESULT 3
ABX15373
ID ABX15373 standard; DNA; 270 BP.
XX
AC ABX15373;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 13-14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to

CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 270;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGGACTCTGGCCCCCAGTGTGA 51
DB 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGGACTCTGGCCCCCAGTGTGA 51

RESULT 4
ABX15374/c
ID ABX15374 standard; DNA; 321 BP.
XX
AC ABX15374;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #2.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct

CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGGACTCTGGCCCCCAGTGTGA 51
DB 102 CATGGGCGCACGGGCACTCCCGTGTCTCTGGACTCTGGCCCCCAGTGTGA 52

RESULT 5
ABX15375/c
ID ABX15375 standard; DNA; 372 BP.
XX
AC ABX15375;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #3.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct


```
FT      /number= 6
FT      7191. .7301
FT      /tag= m
FT      /number= 7
FT      7302. .7452
FT      /tag= n
FT      /number= 8
FT      7453. .7664
FT      /tag= o
FT      /number= 8
FT      7665. .7817
FT      /tag= p
FT      /number= 9
FT      7818. .7910
FT      /tag= q
FT      /number= 9
FT      7911. .8077
FT      /tag= r
FT      /number= 10
FT      8078. .9297
FT      /tag= s
FT      /number= 10
FT      9298. .9479
FT      /tag= t
FT      /number= 11
FT      9480. .10162
FT      /tag= u
FT      /number= 12
FT      10163. .11594
FT      /tag= v
FT      /number= 10
XX
XX      EP787801-A2.
XX
XX      06-AUG-1997.
XX
XX      31-JAN-1997; 97EP-00101622.
XX
XX      01-FEB-1996; 96JP-00016590.
XX
XX      (MITS ) MITSUBISHI CORP.
XX      (MITU ) MITSUBISHI CHEM CORP.
XX
XX      Murase M, Murase J, Hayakawa T, Imamura J, Iwabuchi M;
XX      WPI, 1997-387486/36.
XX
XX      Increasing storage lipid content in seeds and plants - by inhibiting
XX      cytosolic pyruvate kinase.
XX
XX      Claim 9; Page 20-25; 28pp; English.
XX
XX      This DNA encodes a Rattus norvegicus cytosolic pyruvate kinase. The
XX      storage lipid content of a seed is increased by reducing the activity of
XX      endogenous cytosolic pyruvate kinase in the seed. This is applied to
XX      plants which accumulate storage protein and lipid in the embryo,
XX      particularly an oilseed plant such as soya, sunflower, sesame or
XX      CC biosynthesis increases rapeseed. Inhibitory enzymes involved in amino acid
XX      precursor to the chloroplast
XX
XX      Sequence 13011 BP; 3167 A; 3392 C; 3277 G; 3175 T; 0 U; 0 Other;
XX
XX      Query Match      100.0%; Score 51; DB 2; Length 13011;
XX      Best Local Similarity 100.0%; Pred. No. 6.1e-08;
XX      Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 CATTGGCGGACCGGACCTCCGTGCTCTGACTCTGACCTGCCCCCAGTGTGA 51
XX      |||||||||||||||||||||||||||||||||||||||||||||||||||
XX      3021 CATTGGCGGACCGGACCTCCGTGCTCTGACTCTGCCCCCAGTGTGA 3071
XX
XX      RESULT 8
```

```
ABR2448
XX      ID ABR2448 standard; DNA; 13011 BP.
XX
XX      ABR2448;
XX
XX      26-JUN-2003 (first entry)
XX
XX      Toxicity modelling related rat gene SEQ ID No 2150.
XX
XX      Toxic effect; gene expression profile; renal toxicity; toxicity marker;
XX      database; drug screening; toxicity assay; rat; ds.
XX
XX      Rattus norvegicus.
XX
XX      WO200295000-A2.
XX
XX      28-NOV-2002.
XX
XX      22-MAY-2002; 2002MO-US016173.
XX
XX      22-MAY-2001; 2001US-0292335P.
XX      13-JUN-2001; 2001US-0297523P.
XX      19-JUN-2001; 2001US-0296925P.
XX      10-JUL-2001; 2001US-0303807P.
XX      10-JUL-2001; 2001US-0303808P.
XX      10-JUL-2001; 2001US-0303810P.
XX      28-AUG-2001; 2001US-0315047P.
XX      27-SEP-2001; 2001US-0324928P.
XX      22-OCT-2001; 2001US-0330462P.
XX      01-NOV-2001; 2001US-0330867P.
XX      21-NOV-2001; 2001US-0331805P.
XX      06-DEC-2001; 2001US-0336144P.
XX      19-DEC-2001; 2001US-0340873P.
XX      21-FEB-2002; 2002US-0357842P.
XX      21-FEB-2002; 2002US-0357843P.
XX      21-FEB-2002; 2002US-0357844P.
XX      15-MAR-2002; 2002US-0364134P.
XX      08-APR-2002; 2002US-0370144P.
XX      08-APR-2002; 2002US-0370206P.
XX      08-APR-2002; 2002US-0370247P.
XX      17-APR-2002; 2002US-0372794P.
XX      21-APR-2002; 2002US-0371679P.
XX
XX      (GENE-) GENE LOGIC INC.
XX
XX      Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX      WPI; 2003-148464/14.
XX
XX      Predicting at least one toxic effect of a compound, useful for toxicity
XX      PT modeling, comprises preparing a gene expression profile of a tissue or
XX      PT cell sample exposed to the compound, and comparing the gene expression
XX      PT profile to a database.
XX
XX      Example 4; Page; 446pp; English.
XX
XX      The invention relates to a novel method of predicting at least one toxic
XX      CC effect of a compound. The method comprises a gene expression profile of a
XX      CC tissue or cell sample exposed to the compound, and comparing the gene
XX      CC expression profile to a database comprising at least part of the data or
XX      CC information given in the specification. The methods are useful for
XX      CC predicting at least one toxic effect of a compound, predicting the
XX      CC progression of a toxic effect of a compound, predicting the renal
XX      CC toxicity of a compound, or identifying toxicity markers in tissues or
XX      CC cells exposed to known renal toxin. The genes are useful as toxicity
XX      CC markers in drug screening and toxicity assays, in monitoring disease or
XX      CC physiological states, or disease progression. This polynucleotide
XX      CC represents a rat DNA sequence relating to the toxic effect database
XX      CC described in the specification. NOTE: The sequence data for this patent
XX      CC did not form part of the printed specification, but was obtained in
XX      CC electronic format directly from the World Intellectual Property
XX      CC Organization
```

Sequence 13011 BP, 3167 A, 3392 C, 3277 G, 3175 T, 0 U, 0 Other;
Query Match 100.0%; Score 51; DB 10; Length 13011;
Best Local Similarity 100.0%; Pred. No. 6.1e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGGCGACGGGGCACTCCCGTGTTCCTGAGACTCTGAGCCCCAGTGA 51
DB 3021 CATGGGCGACGGGGCACTCCCGTGTTCCTGAGACTCTGAGCCCCAGTGA 3071
RESULT 9
ADP72914
ID ADP72914 standard; DNA; 13011 BP.
XX
AC ADP72914;
XX
DT 26-AUG-2004 (first entry)
XX
DE Renal toxin progression gene marker #1503.
XX
KW ds; toxic effect; gene expression profile; kidney tissue;
KW differential gene expression; toxicity progression; toxicity marker;
KW drug screening; toxicity assay; kidney pathology; nephritis;
KW kidney necrosis; glomerular injury; tubular injury;
KW focal segmental glomerulosclerosis.
XX
OS Rattus norvegicus.
XX
PN WO2004048598-A2.
PD 10-JUN-2004.
XX
PF 24-NOV-2003; 2003WO-US037556.
XX
PR 22-NOV-2002; 2002US-00301856.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
PI Elashoff M;
XX
DR WPI; 2004-460771/43.
XX
PT Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
PS Claim 11; SEQ ID NO 1503; 266pp; English.
XX
XX The invention relates to a method of predicting (the progression of) a
XX toxic effect of a compound by preparing a gene expression profile of a
XX kidney tissue or cell sample exposed to the compound and comparing the
XX gene expression profile to a database, or detecting the level of gene(s)
XX expression in a tissue or cell sample exposed to the compound, where
XX differential gene expression compared to a control indicates a toxic
XX effect (toxicity progression). The method is useful for predicting (the
XX progression of) at least one toxic effect of a compound. The genes are
XX useful as toxicity markers in drug screening and toxicity assays. The
XX methods are useful for predicting the likelihood that a compound or test
XX agent will induce various specific kidney pathologies, such as nephritis,
XX kidney necrosis, glomerular and tubular injury, or focal segmental
XX glomerulosclerosis. The methods are useful for determining the similarity
XX of a toxic response to one or more individual compounds and for
XX predicting or elucidating the potential cellular pathways influenced,
XX induced or modulated by the compound or test agent. The kit is useful for
XX predicting or modelling the toxic response of a test compound, for
XX monitoring the progression of renal disease states, for identifying genes
XX that show promise as new drug targets and for screening known and newly
XX designed drugs. This sequence corresponds to a gene marker used in the
XX method of the invention. (Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic

format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).
XX
Sequence 13011 BP, 3167 A, 3392 C, 3277 G, 3175 T, 0 U, 0 Other;
Query Match 100.0%; Score 51; DB 12; Length 13011;
Best Local Similarity 100.0%; Pred. No. 6.1e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGGCGACGGGGCACTCCCGTGTTCCTGAGACTCTGAGCCCCAGTGA 51
DB 3021 CATGGGCGACGGGGCACTCCCGTGTTCCTGAGACTCTGAGCCCCAGTGA 3071
RESULT 10
ABX15379
ID ABX15379 standard; DNA; 48 BP.
XX
AC ABX15379;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat liver pyruvate kinase glucose response element DNA fragment #1.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ss; IGBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
PN US2002107198-A1.
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THUL P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Disclosure; Page 10; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a fragment of the rat liver pyruvate kinase glucose response
XX element, used in production of the construct of the invention

```
XX Sequence 48 BP; 5 A; 17 C; 17 G; 9 T; 0 U; 0 Other;
SQ
Query Match          94.1%; Score 48; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GGGCGACAGGGGCACTCCCGTGTCTCTGAGACTTG6CCCCCAGTGA 51
    |||
DB 1 GGGCGACAGGGGCACTCCCGTGTCTCTGAGACTTG6CCCCCAGTGA 48

RESULT 11
ABX15380/c
ID ABX15380 standard; DNA; 51 BP.
XX
AC ABX15380;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat liver pyruvate kinase glucose response element DNA fragment #2.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ss; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Disclosure; Page 10; 37pp; English.
XX
SS The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter, and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a fragment of the rat liver pyruvate kinase glucose response
CC element, used in production of the construct of the invention
XX
```

```
SQ Sequence 51 BP; 10 A; 17 C; 18 G; 6 T; 0 U; 0 Other;
SQ
Query Match          94.1%; Score 48; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GGGCGACAGGGGCACTCCCGTGTCTCTGAGACTTG6CCCCCAGTGA 51
    |||
DB 51 GGGCGACAGGGGCACTCCCGTGTCTCTGAGACTTG6CCCCCAGTGA 4

RESULT 12
ADB83565
ID ADB83565 standard; DNA; 30 BP.
XX
AC ADB83565;
XX
DT 29-JAN-2004 (first entry)
XX
DE ChREBP purification oligonucleotide #2.
XX
KW ss; lipogenesis; glycolysis; glucose metabolism;
KW carbohydrate response element binding protein; ChREBP; obesity; diabetes;
KW vascular disease; carbohydrate metabolism; cardiac disease.
XX
OS Synthetic.
XX
PN US2003124590-A1.
XX
PD 03-JUL-2003.
XX
PF 16-OCT-2002; 2002US-00272206.
XX
PR 16-OCT-2001; 2001US-0329834P.
XX
PA (UYED/) UYEDA K.
XX
PI Uyeda K;
XX
DR WPI; 2004-020312/02.
XX
PT Modulation of the expression of DNA molecule that encodes protein
PT involves use of an agent that induces phosphorylation or
PT dephosphorylation of carbohydrate response element binding protein.
XX
PS Example 1; SEQ ID NO 2; 64pp; English.
XX
SS The invention relates to a method of modulation of expression of a DNA
CC molecule (preferably inhibition of lipogenesis or glycolysis) that
CC encodes a protein involved in glucose metabolism or lipogenesis in a cell
CC which comprises contacting the cell with an agent that induces
CC phosphorylation or dephosphorylation of carbohydrate response element
CC binding protein (ChREBP). The method is useful for treating obesity,
CC diabetes or vascular disease and modulating carbohydrate metabolism, for
CC inhibiting lipogenesis or glycolysis and modulating expression of a DNA
CC molecule that encodes a protein involved in glucose metabolism or
CC lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
CC disease. The present sequence represents an oligonucleotide used in the
CC purification of ChREBP
XX
SQ Sequence 30 BP; 3 A; 11 C; 11 G; 5 T; 0 U; 0 Other;
SQ
Query Match          58.8%; Score 30; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CATGGCGCAGGGGCACTCCCGTGTCC 30
    |||
DB 1 CATGGCGCAGGGGCACTCCCGTGTCC 30

RESULT 13
ADB83567/c
```

```

ID ADE83567 standard; DNA; 27 BP.
XX
XX ADE83567;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX L-pyruvate kinase carboxylate response element #2.
DE
XX
XX ds; lipogenesis; glycolysis; glucose metabolism;
KM carboxylate response element binding protein; ChREBP; obesity; diabetes;
KM vascular disease; carbohydrate metabolism; cardiac disease;
KM L-pyruvate kinase.
XX
XX Unidentified.
OS
XX
XX US2003124590-A1.
PN
XX
XX 03-JUL-2003.
PD
XX
XX 16-OCT-2002; 2002US-00272206.
PF
XX
XX 16-OCT-2001; 2001US-0329834P.
PR
XX
XX (UYED/) UYEDA K.
PA
XX
XX Uyeda K;
PI
XX
XX MPI; 2004-020312/02.
DR
XX
XX Modulation of the expression of DNA molecule that encodes protein
PT involves use of an agent that induces phosphorylation or
PT dephosphorylation of carbohydrate response element binding protein.
XX
XX Example 2; SEQ ID NO 4; 64pp; English.
PS
XX
XX The invention relates to a method of modulation of expression of a DNA
CC molecule (preferably inhibition of lipogenesis or glycolysis) that
CC encodes a protein involved in glucose metabolism or lipogenesis in a cell
CC which comprises contacting the cell with an agent that induces
CC phosphorylation or dephosphorylation of carbohydrate response element
CC binding protein (ChREBP). The method is useful for treating obesity,
CC diabetes or vascular disease and modulating carbohydrate metabolism, for
CC inhibiting lipogenesis or glycolysis and modulating expression of a DNA
CC molecule that encodes a protein involved in glucose metabolism or
CC lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
CC disease. The present sequence represents a L-pyruvate kinase carboxylate
CC response element.
XX
XX Sequence 27 BP; 4 A; 11 C; 10 G; 2 T; 0 U; 0 Other;
SQ
XX
XX Query Match 52.9%; Score 27; DB 12; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 7;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GGGCGCAGCGGCACTCCCGTGTTC 30
DB 27 GGGCGCAGCGGCACTCCCGTGTTC 1

```

```

XX
XX Unidentified.
OS
XX
XX US2003124590-A1.
PN
XX
XX 03-JUL-2003.
PD
XX
XX 16-OCT-2002; 2002US-00272206.
PF
XX
XX 16-OCT-2001; 2001US-0329834P.
PR
XX
XX (UYED/) UYEDA K.
PA
XX
XX Uyeda K;
PI
XX
XX MPI; 2004-020312/02.
DR
XX
XX Modulation of the expression of DNA molecule that encodes protein
PT involves use of an agent that induces phosphorylation or
PT dephosphorylation of carbohydrate response element binding protein.
XX
XX Example 2; SEQ ID NO 3; 64pp; English.
PS
XX
XX The invention relates to a method of modulation of expression of a DNA
CC molecule (preferably inhibition of lipogenesis or glycolysis) that
CC encodes a protein involved in glucose metabolism or lipogenesis in a cell
CC which comprises contacting the cell with an agent that induces
CC phosphorylation or dephosphorylation of carbohydrate response element
CC binding protein (ChREBP). The method is useful for treating obesity,
CC diabetes or vascular disease and modulating carbohydrate metabolism, for
CC inhibiting lipogenesis or glycolysis and modulating expression of a DNA
CC molecule that encodes a protein involved in glucose metabolism or
CC lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
CC disease. The present sequence represents a L-pyruvate kinase carboxylate
CC response element.
XX
XX Sequence 27 BP; 2 A; 10 C; 11 G; 4 T; 0 U; 0 Other;
SQ
XX
XX Query Match 52.9%; Score 27; DB 12; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 7;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 GGGCGCAGCGGCACTCCCGTGTTC 30
DB 1 GGGCGCAGCGGCACTCCCGTGTTC 27

```

```

RESULT 14
ADE83566
ID ADE83566 standard; DNA; 27 BP.
XX
XX ADE83566;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX L-pyruvate kinase carboxylate response element #1.
DE
XX
XX ds; lipogenesis; glycolysis; glucose metabolism;
KM carboxylate response element binding protein; ChREBP; obesity; diabetes;
KM vascular disease; carbohydrate metabolism; cardiac disease;
KM L-pyruvate kinase.
XX

```

```

RESULT 15
ADE83614
ID ADE83614 standard; DNA; 26 BP.
XX
XX ADE83614;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX ChREBP gel shift assay gene sequence #1.
DE
XX
XX ss; lipogenesis; glycolysis; glucose metabolism;
KM carboxylate response element binding protein; ChREBP; obesity; diabetes;
KM vascular disease; carbohydrate metabolism; cardiac disease;
KM gel shift assay.
XX
XX Unidentified.
OS
XX
XX US2003124590-A1.
PN
XX
XX 03-JUL-2003.
PD
XX
XX 16-OCT-2002; 2002US-00272206.
PF
XX
XX 16-OCT-2001; 2001US-0329834P.
PR
XX
XX (UYED/) UYEDA K.
PA

```

XX Uyeda K;
PI
XX

DR WPI; 2004-020312/02.

XX Modulation of the expression of DNA molecule that encodes protein
PT involves use of an agent that induces phosphorylation or
PT dephosphorylation of carbohydrate response element binding protein.

XX Example 14; SEQ ID NO 51; 64pp; English.

PS
XX The invention relates to a method of modulation of expression of a DNA
CC molecule (preferably inhibition of lipogenesis or glycolysis) that
CC encodes a protein involved in glucose metabolism or lipogenesis in a cell
CC which comprises contacting the cell with an agent that induces
CC phosphorylation or dephosphorylation of carbohydrate response element
CC binding protein (ChREBP). The method is useful for treating obesity, for
CC diabetes or vascular disease and modulating carbohydrate metabolism, for
CC inhibiting lipogenesis or glycolysis and modulating expression of a DNA
CC molecule that encodes a protein involved in glucose metabolism or
CC lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
CC disease. The present sequence represents a ChREBP gel shift assay gene
CC sequence.
XX

SO Sequence 26 BP; 2 A; 9 C; 11 G; 4 T; 0 U; 0 Other;

Query Match 51.0%; Score 26; DB 12; Length 26;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGCGGCACTCCCGTGATC 29
|||||
DB 1 GGGCGACGCGGCACTCCCGTGATC 26

Search completed: September 1, 2005, 00:34:35
Job time : 60.0797 secs

This Page Blank (uspto)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 : Search time 148.534 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51

Sequence: 1 catggcgccgacggcgacac.....ggactctgccccagctgta 51

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 733684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	51	9	US-09-972-916A-1
2	51	100.0	270	9	US-09-972-916A-3
3	51	100.0	321	9	US-09-972-916A-4
4	51	100.0	372	9	US-09-972-916A-5
5	51	100.0	423	9	US-09-972-916A-6
6	51	100.0	13011	17	US-10-388-934-36
7	51	100.0	13011	18	US-10-152-319A-2150

C	8	36.6	71.8	31703	17	US-10-085-117-172	Sequence 172, App
9	27	58.9	30	15	US-10-272-206-2	Sequence 2, Appl	
10	27	52.9	27	15	US-10-272-206-3	Sequence 3, Appl	
11	27	52.9	27	15	US-10-272-206-4	Sequence 4, Appl	
12	26	51.0	26	15	US-10-272-206-5	Sequence 5, Appl	
13	24.6	48.2	55001	17	US-10-160-497-4	Sequence 4, Appl	
14	24.6	48.2	55001	17	US-10-348-750-4	Sequence 4, Appl	
15	24.6	48.2	55001	21	US-10-991-147-4	Sequence 4, Appl	
16	24.4	47.8	3138	9	US-09-880-107-1116	Sequence 1716, Ap	
17	24.4	47.8	3138	17	US-10-172-118-909	Sequence 909, App	
18	24.4	47.8	3138	18	US-10-342-887-909	Sequence 909, App	
19	24.4	47.8	3138	20	US-10-723-860-176	Sequence 176, App	
20	24.4	47.8	3138	22	US-10-756-149-213	Sequence 213, App	
21	24.4	47.8	3511	20	US-10-723-860-4919	Sequence 4919, App	
22	24	47.1	209	16	US-10-029-386-26227	Sequence 26227, App	
23	24	47.1	525	16	US-10-029-386-12527	Sequence 12527, A	
24	24	47.1	1188	14	US-10-094-417-3	Sequence 3, Appl	
25	24	47.1	1188	24	US-11-086-846-3	Sequence 3, Appl	
26	24	47.1	1191	18	US-10-343-650A-45	Sequence 45, Appl	
27	24	47.1	1311	15	US-10-029-426-1	Sequence 1, Appl	
28	24	47.1	1333	15	US-10-017-161-1921	Sequence 1921, Ap	
29	24	47.1	1498	9	US-09-826-508-37	Sequence 37, Appl	
30	24	47.1	1546	15	US-10-225-567A-584	Sequence 584, App	
31	24	47.1	1579	14	US-10-073-885-11	Sequence 11, Appl	
32	24	47.1	1595	9	US-09-798-710-1	Sequence 1, Appl	
33	24	47.1	1905	21	US-10-505-486-226	Sequence 226, App	
34	24	47.1	2046	9	US-09-764-893-130	Sequence 130, App	
35	24	47.1	2046	9	US-09-764-886-84	Sequence 84, Appl	
36	24	47.1	2046	9	US-09-764-886-84	Sequence 875, App	
37	24	47.1	2046	10	US-09-989-442-155	Sequence 155, App	
38	24	47.1	2046	10	US-09-764-886-84	Sequence 84, Appl	
39	24	47.1	2046	11	US-09-764-886-84	Sequence 1209, App	
40	24	47.1	2046	14	US-10-073-885-130	Sequence 130, App	
41	24	47.1	2046	14	US-10-103-313-592	Sequence 592, App	
42	23.8	46.7	27	15	US-10-272-206-5	Sequence 5, Appl	
43	23.8	46.7	27	15	US-10-272-206-6	Sequence 6, Appl	
44	23.8	46.7	27	15	US-10-272-206-6	Sequence 7, Appl	
45	23.8	46.7	27	15	US-10-272-206-8	Sequence 8, Appl	

ALIGNMENTS

RESULT 1
US-09-972-916A-1
; Sequence 1, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972_916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; US-09-972-916A-1

Query Match 100.0%; Score 51; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 8.1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 CATGGCGCAGGGGACATCCCGTCTCTGAGCTTGGCCCCAGGTGA 51
DB 1 CATGGCGCAGGGGACATCCCGTCTCTGAGCTTGGCCCCAGGTGA 51
RESULT 2

```
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 270;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
Db
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
```

```
RESULT 3
US-09-972-916A-4/c
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 321;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
Db
102 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 52
```

```
RESULT 4
US-09-972-916A-5/c
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
```

```
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 372;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
Db
153 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 103
```

```
RESULT 5
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
```

```
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 423;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
Db
1 CATGGGCGACGCGGCACTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
```

```
RESULT 6
US-10-388-934-36
; Sequence 36, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 13011
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-36
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 17; Length 13011;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 CATGGCGCAGCGGCACTCCGTTGTTCTGACTGTGCCCCCAGTGA 51
Db 3021 CATGGCGCAGCGGCGCACTCCGTTGTTCTGACTGTGCCCCCAGTGA 3071

RESULT 7
US-10-152-319A-2150
Sequence 2150, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higge, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2150
LENGTH: 13011
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. X05684
US-10-152-319A-2150

Query Match 100.0%; Score 51; DB 18; Length 13011;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGCACTCCGTTGTTCTGACTGTGCCCCCAGTGA 51
Db 3021 CATGGCGCAGCGGCGCACTCCGTTGTTCTGACTGTGCCCCCAGTGA 3071

RESULT 8
US-10-085-117-172/C
Sequence 172, Application US/10085117
Publication No. US2003023334A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586

PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 172
LENGTH: 31703
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variation
LOCATION: (1)...(31703)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-172

Query Match 71.8%; Score 36.6; DB 17; Length 31703;
Best Local Similarity 90.7%; Pred. No. 0.00041;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACCTCCGTTGTTCTGACTGTGCCCCCAGTGA 51
Db 2120 CACGGGCGACCTCCGTTGTTCTGACTGTGCCCCCAGTGA 2078

RESULT 9
US-10-272-206-2
Sequence 2, Application US/102722206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495 0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-10-272-206-2

Query Match 58.8%; Score 30; DB 15; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGCACTCCGTTGTTCC 30
Db 1 CATGGCGCAGCGGCGCACTCCGTTGTTCC 30

RESULT 10
US-10-272-206-3
Sequence 3, Application US/102722206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495 0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Oligonucleotide from wild type pyruvate kinase
OTHER INFORMATION: carbohydrate response element (CHRE)
US-10-272-206-3

Query Match 52.9%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGTTC 30
DB 1 GGGCGACGGGGCACTCCCGTGTTC 27

RESULT 11
US-10-272-206-4/C

Sequence 4, Application US/10272206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495.0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide from wild type pyruvate kinase
OTHER INFORMATION: carbohydrate response element (CHRE)
US-10-272-206-4

Query Match 52.9%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGTTC 30
DB 27 GGGCGACGGGGCACTCCCGTGTTC 1

RESULT 12
US-10-272-206-51

Sequence 51, Application US/10272206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495.0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 51
LENGTH: 26
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-10-272-206-51

Query Match 51.0%; Score 26; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGTTC 29
DB 1 GGGCGACGGGGCACTCCCGTGTTC 26

RESULT 13
US-10-160-497-4

Sequence 4, Application US/10160497
Publication No. US20030224513A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF NOTCH1 EXPRESSION
FILE REFERENCE: RTS-0386
CURRENT APPLICATION NUMBER: US/10/160,497
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 145
SEQ ID NO 4
LENGTH: 55001
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-160-497-4

Query Match 48.2%; Score 24.6; DB 17; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGTTCGACTGCGCCCACTGT 50
DB 27032 GGGTCACGGGGCTCTCGCTGCTTCCGCCCTCTCCCACTGT 27078

RESULT 14
US-10-348-750-4

Sequence 4, Application US/10348750
Publication No. US20030225019A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: NOTCH1 INHIBITORS FOR INDUCING APOPTOSIS
FILE REFERENCE: ISPH-0729
CURRENT APPLICATION NUMBER: US/10/348,750
CURRENT FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: 10/160,497
NUMBER OF SEQ ID NOS: 146
SEQ ID NO 4
LENGTH: 55001
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-348-750-4

Query Match 48.2%; Score 24.6; DB 17; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGTTCGACTGCGCCCACTGT 50
DB 27032 GGGTCACGGGGCTCTCGCTGCTTCCGCCCTCTCCCACTGT 27078

RESULT 15
US-10-991-147-4

Sequence 4, Application US/10991147
Publication No. US20050096292A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Dobie

```
; APPLICANT: Brich Koller
; TITLE OF INVENTION: NORCH1 INHIBITORS FOR INDUCING APOPTOSIS
; FILE REFERENCE: ISPH-0729
; CURRENT APPLICATION NUMBER: US/10/991,147
; CURRENT FILING DATE: 2004-11-17
; PRIOR APPLICATION NUMBER: US/10/348,750
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 10/160,497
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 146
; SEQ ID NO 4
; LENGTH: 55001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-991-147-4
```

```
Query Match 48.2%; Score 24.6; DB 21; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5; Mismatches 14; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
Qy 4 GGGCGCAGCGGGGCACTCCCGTGGTTCCTGGAAGCTGAGCCCACTGT 50
Db 27032 GGGTCCACGGGGCTCGCTGGGCTCCGCGCCCTCTCCCCCACTGTGT 27078
```

```
Search completed: September 1, 2005, 16:44:00
Job time : 164.534 secs
```

This Page Blank (uspto)

Db	943	CCACTGCCCGCCGAGACACAAACCAGGAGCATTTGAACACTTGACACGCGCATCTGCCC	1002		
Qy	191	AGAGAGCTGTGACCAACCACTTCCTCGCTACTAGCTA	214		
Db	1003	AGAGAGCTGTGACCAACCACTTCCTCGCTACTAGCTA	1036		
RESULT 2	AC136382	185148 bp	DNA	linear	HTG 01-NOV-2002
LOCUS	Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS				
DEFINITION	*** 63 unordered pieces.				
ACCESSION	AC136382	1	GI:24462257		
VERSION	AC136382.1				
KEYWORDS	HTG; PHASE1.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 185148)				
AUTHORS	Munzy,D.Marie., Merker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alb Brooks,S., Amin,A., Angiano,D., Aryalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barneshead,M., Benahmed,F., Biswalto,K., Blair,J., Blankemburg,K., Blyth,P., Brown,M., Bryant,N., Bubay,C., Blancheburg,K., Burrell,K., Calderon,E., Cadenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,M., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,T., Fan,G., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gatta,R., Garcia,A., Garner,T., Garza,M., Georgesorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevarra,W., Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Kapachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C., Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lourenshewa,L., Louisgeed,H., Locrado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindratne,M., Mahmood,M., Malloy,K., Mangum,A., Mangum,B., Nagua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., Mcleod,M., Mcneill,T., Meenen,E., Mlivosavljevic,A., Miller,G., Mirja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L., Narkantis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwokilemeh,O., Okunnu,G., Olarnpneegoon,A., Pal,S., Paks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Plankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-D., Puzo,M., Quitoz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Steele,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Sytek,A., Tabors,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmari,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,T., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,J., Wleciyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,D., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 185148)				
AUTHORS	Rat Genome Sequencing Consortium.				

TITLE Direct Submission
 JOURNAL Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KDG5
 Center clone name: CH230-97018
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 135613 bases at least Q40
 Consensus quality: 140849 bases at least Q30
 Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 63 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	1010:	contig of 1010 bp in length
*	1011	1110: gap of unknown length
*	1111	2608: contig of 1498 bp in length
*	2609	2708: gap of unknown length
*	2709	3915: contig of 1207 bp in length
*	3916	4015: gap of unknown length
*	4016	5547: contig of 1532 bp in length
*	5548	5647: gap of unknown length
*	5648	7119: contig of 1472 bp in length
*	7120	7219: gap of unknown length
*	7220	8885: contig of 1666 bp in length
*	8886	8985: gap of unknown length
*	8986	10258: contig of 1273 bp in length
*	10259	10358: gap of unknown length
*	10359	11626: contig of 1268 bp in length
*	11627	11726: gap of unknown length
*	11727	13488: contig of 1762 bp in length
*	13489	13588: gap of unknown length
*	13589	15122: contig of 1534 bp in length
*	15123	15222: gap of unknown length
*	15223	16946: contig of 1724 bp in length
*	16947	17046: gap of unknown length
*	17047	18164: contig of 1118 bp in length
*	18165	18264: gap of unknown length
*	18265	19678: contig of 1414 bp in length
*	19679	19778: gap of unknown length
*	19779	21687: contig of 1909 bp in length
*	21688	21787: gap of unknown length
*	21788	22892: contig of 1105 bp in length
*	22893	22992: gap of unknown length
*	22993	24336: contig of 1344 bp in length
*	24337	24436: gap of unknown length
*	24437	26617: contig of 2181 bp in length
*	26618	26717: gap of unknown length
*	26718	28577: contig of 1860 bp in length
*	28578	28677: gap of unknown length
*	28678	30823: contig of 2146 bp in length
*	30824	30923: gap of unknown length
*	30924	32266: contig of 1343 bp in length
*	32267	32366: gap of unknown length
*	32367	33894: contig of 1528 bp in length
*	33895	33994: gap of unknown length
*	33995	35373: contig of 1379 bp in length


```

* 35374 35473: gap of unknown length
* 35474 37295: contig of 1822 bp in length
* 37296 37395: gap of unknown length
* 37396 38420: contig of 1025 bp in length
* 38421 38520: gap of unknown length
* 38521 39974: contig of 1454 bp in length
* 39975 40074: gap of unknown length
* 40075 42283: contig of 2209 bp in length
* 42284 42383: gap of unknown length
* 42384 43721: contig of 1338 bp in length
* 43722 43821: gap of unknown length
* 43822 46538: contig of 2717 bp in length
* 46539 46638: gap of unknown length
* 46639 48621: contig of 1983 bp in length
* 48622 48721: gap of unknown length
* 48722 50602: contig of 1881 bp in length
* 50603 50702: gap of unknown length
* 50703 53841: contig of 3139 bp in length
* 53842 53941: gap of unknown length
* 53942 56195: contig of 2254 bp in length
* 56196 56295: gap of unknown length
* 56296 58949: contig of 2654 bp in length
* 58950 59049: gap of unknown length
* 59050 61032: contig of 1983 bp in length
* 61033 61132: gap of unknown length
* 61133 62942: contig of 1810 bp in length
* 62943 63042: gap of unknown length
* 63043 66045: contig of 3003 bp in length
* 66046 66145: gap of unknown length
* 66146 68197: contig of 2052 bp in length
* 68198 68297: gap of unknown length
* 68298 72235: contig of 3938 bp in length
* 72236 72335: gap of unknown length
* 72336 75803: contig of 3468 bp in length
* 75804 75903: gap of unknown length
* 75904 78355: contig of 2452 bp in length
* 78356 78455: gap of unknown length
* 78456 81458: contig of 3003 bp in length
* 81459 81558: gap of unknown length
* 81559 84673: contig of 3115 bp in length
* 84674 84773: gap of unknown length
* 84774 87739: contig of 2966 bp in length
* 87740 87839: gap of unknown length
* 87840 91535: contig of 3696 bp in length
* 91536 91635: gap of unknown length
* 91636 95522: contig of 3887 bp in length
* 95523 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98830: gap of unknown length
* 98831 101618: contig of 2768 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117630: contig of 3966 bp in length
* 117631 117731: gap of unknown length
* 117732 122502: contig of 4772 bp in length
* 122503 122602: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126645 126744: gap of unknown length

```

Query Match 97.0%; Score 212.4; DB 2; Length 185148;
 Best Local Similarity 99.5%; Pred. No. 2.4e-53;
 Matches 213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTTATTTTGAACGCGGAGATCTTGAACAAGCTGCTCCCTGACATCA 60
 DB 99253 TCACAGCAAAACAACTTTATTTTGAACGCGGAGATCTTGAACAAGCTGCTCCCTGACATCA 99312
 QY 61 TTAACCGGTGCTGCGAGCCAGCCCTTATTAAGGCGCTGGATAGGCCAGCCAGCATGCT 120

```

DB 99313 TTAACCGGTGCTGCGAGCCAGCCCTTATTAAGGCGCTGGATAGGCCAGCCAGCATGCT 99372
QY 121 CCACCTGCCCGCCGAGACAAACCCAGCAGCATTTGACATGCAACGCGCCATCTGCC 180
DB 99373 CCACCTGCCCGCCGAGACAAACCCAGCAGCATTTGACATGCAACGCGCCATCTGCC 99432
QY 181 AAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 214
DB 99433 AAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 99466

RESULT 3
MTLGF
LOCUS MTLGF 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1363)
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
Hepatology 19 (3), 656-665 (1994)
JOURNAL MEDLINE 94164648
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohn,K.L., Waddie,J.R. and Taub,R.
Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
Nucleic Acids Res.
3 (bases 1 to 1363)
REFERENCE Taub,R.A.
AUTHORS Direct Submision
JOURNAL Submitted (23-JUN-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/feature="AP-2 consensus site"
692..706
/feature="insulin-responsive element"
720..741
/feature="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
/feature="TATA_signal"
792..1336
/feature="exon"
792
/feature="misc_feature"
964..1336
/feature="CDS"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

```

ORIGIN

/translation="MPEFLTVVSWPFLILLSFQIGVAGAPQPMHCAPCTARLGLCP
PVPASCPBESRPAGCGCCCTCCALPMGACAGVATPACAGLSRALPGEBRPLHALTRG
QGASLPEPAPATSTLIFSQHE"

Query Match 86.0%; Score 188.4; DB 10; Length 1363;
Best Local Similarity 92.5%; Pred. No. 4.1e-46;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTGTAACACGGGATCTTACAGCGTCCCTGCAATCA 60
DB 677 TCACAGCAAAACAACTTATTGTAACACGGGATCTTACAGCGTCCCTGCAATCA 736
QY 61 TTAACCCGTGCTGCGAGACACGCCCTTCATAGAGCCCTGATGAGCCAGCATGCT 120
DB 737 TTAACCTGTGCGGACACAGCCAGCCCTTCATAGAGCTGTGGTATGACACGACGATGCT 796
QY 121 CCACGTGCGCGGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATTTGCC 180
DB 797 CCACGTGCGCGGAGACACACACCCAGGAGCATTTGAACACTGCACACGCGCATTTGCC 856
QY 181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 214
DB 857 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 890

RESULT 4
LOCUS AL607124 192843 bp DNA linear ROD 11-APR-2002
DEFINITION Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
ACCESSION AL607124
VERSION AL607124 GI:20145926
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Oliver K.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonesrequest@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is
from the RPCT-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES

source

1..192843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

ORIGIN

/clone="RP23-20C9"
/clone_1ib="RPCT-23"

Query Match 86.0%; Score 188.4; DB 10; Length 192843;
Best Local Similarity 92.5%; Pred. No. 4.5e-46;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTGTAACACGGGATCTTACAGCGTCCCTGCAATCA 60
DB 2429 TCACAGCAAAACAACTTATTGTAACACGGGATCTTACAGCGTCCCTGCAATCA 2488
QY 61 TTAACCCGTGCTGCGAGACACGCCCTTCATAGAGCCCTGATGAGCCAGCATGCT 120
DB 2489 TTAACCTGTGCGGACACAGCCAGCCCTTCATAGAGCTGTGGTATGACACGACGATGCT 2548
QY 121 CCACGTGCGCGGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATTTGCC 180
DB 2549 CCACGTGCGCGGAGACACACACCCAGGAGCATTTGAACACTGCACACGCGCATTTGCC 2608
QY 181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 214
DB 2609 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 2642

RESULT 5
LOCUS BC013345 1569 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus insulin-like growth factor binding protein 1, mRNA
(cDNA clone MGC:14075 IMAGE:416189), complete cds.
ACCESSION BC013345
VERSION BC013345.1 GI:15426482
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietzenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheerz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Millamy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Guarnatone, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huily, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Heltan, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Girmwood, J., Schmutz, V., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Merris, M.A.

TITLE human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCES 2 (bases 1 to 1569)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: http://image.llnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source Location/Qualifiers

1..1569
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:14075 IMAGE:4161889"
/cissue_type="Liver, normal. 5 month old male mouse."
/clone_1ib="NCI_CGAP_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocusID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="Insulin-like growth factor binding protein 1"
/protein_id="AAH13345.1"
/db_xref="GI:15426483"
/db_xref="LocusID:16006"
/db_xref="MGI:96436"
/translation="MPEFTVYSWPELILTSQIVGAAGAPQPHWCAPCTARLGLCPYPASPEISRPAGCGCCCTCALPMGAAGCAVATACAOGLSCRAUGEPRLHALTRGOGACVPEPAAPATSTLSSQHEAKAAVAVSADLESPEMTEQLDSFHLMAFREDQPLMNAISTVSMARBEIADLKWKEPCORLYLVLELAAQOKAGDELYKPYLPNCNNGFYSKQCEITSDGBAGLCWCVCYPMGSKKIPGSLETRGDPNCHQYFVNH"

CDS

1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocusID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="Insulin-like growth factor binding protein 1"
/protein_id="AAH13345.1"
/db_xref="GI:15426483"
/db_xref="LocusID:16006"
/db_xref="MGI:96436"
/translation="MPEFTVYSWPELILTSQIVGAAGAPQPHWCAPCTARLGLCPYPASPEISRPAGCGCCCTCALPMGAAGCAVATACAOGLSCRAUGEPRLHALTRGOGACVPEPAAPATSTLSSQHEAKAAVAVSADLESPEMTEQLDSFHLMAFREDQPLMNAISTVSMARBEIADLKWKEPCORLYLVLELAAQOKAGDELYKPYLPNCNNGFYSKQCEITSDGBAGLCWCVCYPMGSKKIPGSLETRGDPNCHQYFVNH"

ORIGIN

Query Match 44.8%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 9e-19;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 104 TGGCCAGCAGCATGTCTCCAGTCCGCGCGAGACACAAACCCAGGAGCATTTGAACACTG 163
DB 20 TGGGCGCCAGCATGTCTCCAGTCCGCGCGAGACACACCCAGGAGCATTTGAACACTG 79
QY 164 CACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
DB 80 CACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 130

RESULT 6
AX401932 1500 bp DNA linear PAT 06-JUN-2002
LOCUS Sequence 1608 from Patent WO0210453.
DEFINITION AX401932
ACCESSION AX401932
VERSION AX401932.1 GI:21338112
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and Elashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;

Gene Logic, Inc. (US)
Location/Qualifiers

source 1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTTCGCTACTAGCTA 214
DB 68 CACTTCGCTACTAGCTA 85

RESULT 7
AX827271 1500 bp DNA linear PAT 12-DEC-2003
LOCUS Sequence 5 from Patent EP1344834.
DEFINITION AX827271
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
F. HOFMANN-LA ROCHE AG (CH)
FEATURES
source Location/Qualifiers
1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTTCGCTACTAGCTA 214
DB 68 CACTTCGCTACTAGCTA 85

RESULT 8

LOCUS RATTGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (RIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Mohm, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.

TITLE The gene encoding rat insulinlike growth factor-binding protein 1 is rapidly and highly induced in regenerating liver

JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)

MEDLINE 9141487

PUBMED 1705004

COMMENT Original source text: Rat, cDNA to mRNA.

FEATURES

source

1..1500

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Fisher"

/db_xref="taxon:10116"

/tissue_type="regenerating liver"

/dev_stage="adult"

160..978

/codon_start=1

/product="IGF binding protein-1"

/protein_id="AA41380.1"

/db_xref="GI:204733"

/translation="MPEFLTVVSWPFLILSFQVRVVGAPQPMHCACTAERLELCP PVPASCEISRPAAGCGCCPTCALPIGACGVATPAACQGLSCRALGPEPRPLHALTRG OGACVLEPAAPATSSLSGSHBEAKAVASDELAESPMTESQLDSFHLMAPSRED QPIIMNAISTYSSMRAREITDIKKMKEPCCRELKYLERLAQAQKAGDEITYKFLPN CNKNGFYHSKQCEISLDGBAGLCWCVPWMSGKKIPGSLFTRGDPNCHQYFNVO"

160..234

235..975

/product="IGF binding protein-1"

ORIGIN

Query Match 34.9%; Score 76.4; DB 10; Length 1500;

Best Local Similarity 98.7%; Pred. No. 3.7e-12;

Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCACAGCGCCATCTGCCAGAGAGCTGTGACAC 196

|||||

DB 8 CACAAACCCAGGAGCATTTGAACACTGCACAGCGCCATCTGCCAGAGAGCTGTGACAC 67

QY 197 CACTTCGGCTACTAGCTA 214

|||||

DB 68 CACTTCGGCTACTATCTA 85

RESULT 9

AX163782

LOCUS Sequence 46 from Patent W00138579. 5001 bp DNA linear PAT 22-JUN-2001

DEFINITION AX163782

ACCESSION AX163782.1 GI:14544878

VERSION

KEYWORDS

SOURCE

ORGANISM Rattus norvegicus (Norway rat)

Rattus norvegicus

Bufo variegatus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 Gould-Rothberg,B.E., Dipippo,V.A., Ramsen,T.M. and Gerwein,R.W. Method of identifying toxic agents using nasal-induced differential gene expression in liver

JOURNAL Patent: WO 0138579-A 46 31-MAY-2001;

Curagen Corporation (US)

FEATURES

source

1..5001

/organism="Rattus norvegicus"

/mol_type="unassigned DNA"

/db_xref="taxon:10116"

ORIGIN

Query Match 34.9%; Score 76.4; DB 6; Length 5001;

Best Local Similarity 98.7%; Pred. No. 3.7e-12;

Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCACAGCGCCATCTGCCAGAGAGCTGTGACAC 196

|||||

DB 1 CACAAACCCAGGAGCATTTGAACACTGCACAGCGCCATCTGCCAGAGAGCTGTGACAC 60

QY 197 CACTTCGGCTACTAGCTA 214

|||||

DB 61 CACTTCGGCTACTATCTA 78

RESULT 10

RATIGFBA

LOCUS 5001 bp DNA linear ROD 30-NOV-1995

DEFINITION Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene, complete cds.

ACCESSION L22979

VERSION L22979.1 GI:1098472

KEYWORDS insulin-like growth factor binding protein-1.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Bufo variegatus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 5001)

Lacson,R., Oehler,D., Yang,E., Goswami,R. and Unterman,T. Dideoxy sequencing and structural analysis of the rat insulin-like growth factor binding protein-1 gene

JOURNAL Biochim. Biophys. Acta 1218 (1), 95-98 (1994)

MEDLINE 94250701

PUBMED 7514892

COMMENT On Nov 30, 1995 this sequence version replaced gi:385167. Original source text: Rattus norvegicus (strain Sprague-Dawley) DNA.

FEATURES

source

1..5001

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/sex="male"

/tissue_type="liver"

/dev_stage="adult"

/tissue_lib="Clontech EMBL-3 SP6/T7"

60..525

/number=1

60..152

153..4221

/gene="IGFBP-1"

join(153..525,1850..2034,2874..3002,4090..4221)

/gene="IGFBP-1"

/codon_start=1

/product="insulin-like growth factor binding protein"

/protein_id="AA82581.1"

/db_xref="GI:1098473"

/translation="MPEFLTVVSWPFLILSFQVRVVGAPQPMHCACTAERLELCP PVPASCEISRPAAGCGCCPTCALPIGACGVATPAACQGLSCRALGPEPRPLHALTRG OGACVLEPAAPATSSLSGSHBEAKAVASDELAESPMTESQLDSFHLMAPSRED QPIIMNAISTYSSMRAREITDIKKMKEPCCRELKYLERLAQAQKAGDEITYKFLPN CNKNGFYHSKQCEISLDGBAGLCWCVPWMSGKKIPGSLFTRGDPNCHQYFNVO"

526..1849

/gene="IGFBP-1"

/number=1

1850..2034

/gene="IGFBP-1"

/number=2

2035..2873

/gene="IGFBP-1"

/number=2

2874..3002

/gene="IGFBP-1"

/number=3

3003..4089

/gene="IGFBP-1"

/number=3

4090..4743

/number=4

exon

intron

exon

intron

exon

```

ORIGIN      3'UTR      4222. .4743
Query Match 34.9%; Score 76.4; DB 10; Length 5001;
Best Local Similarity 98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCGAGAGATTGAACACTGCACAGCGGCATCGCCACAGAGCTGTGACAC 196
DB 1 CACAAACCCGAGAGATTGAACACTGCACAGCGGCATCGCCACAGAGCTGTGACAC 60

QY 197 CACTTCGGCTACTAGCTA 214
DB 61 CACTTCGGCTACTAGCTA 78

RESULT 11
LOCUS      AY560836      1507 bp      mRNA      linear      ROD 22-MAR-2004
DEFINITION Spermophilus tridecemlineatus insulin-like growth factor binding
ACCESSION  AY560836
VERSION     AY560836.1 GI:45505308
KEYWORDS
SOURCE
ORGANISM   Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
            Spermophilus.
REFERENCE  1 (bases 1 to 1507)
            Li,Y., Klimenta,D. and Hallenbeck,J.M.
            Cloning and characterization of insulin-like growth factor binding
            protein 1 (IGFBP-1) from thirteen-lined ground squirrel
            Unpublished
JOURNAL    2 (bases 1 to 1507)
AUTHORS    Li,Y., Klimenta,D. and Hallenbeck,J.M.
TITLE      Direct Submission
JOURNAL    Submitted (27-FEB-2004) Strokebranch, NINDS/NIH, 36 Convent Dr.,
            Bethesda, MD 20892, USA
FEATURES   Location/Qualifiers
            source          1..1507
                        /organism="Spermophilus tridecemlineatus"
                        /mol_type="mRNA"
                        /db_xref="taxon:43179"
            gene            1..1507
                        /gene="IGFBP1"
            CDS             689..1507
                        /gene="IGFBP1"
                        /codon_start=1
                        /product="insulin-like growth factor binding protein 1"
                        /protein_id="AA567029.1"
                        /db_xref="GI:45505309"
ORIGIN
Query Match 34.8%; Score 76.2; DB 10; Length 1507;
Best Local Similarity 67.8%; Pred. No. 4.2e-12;
Matches 143; Conservative 0; Mismatches 53; Indels 15; Gaps 2;

QY 2 CACAGCAAAACAACTATTATTTGAACACGGGATCTTACACGCTGCTGACATCAT 61
DB 431 CAGAGCAAAACAACTATTATTTGAACACGGGCTCTTACACGCGCGCTGGCATCAT 490
QY 62 TAAACC-----GTGCTGCCAGGCAAGCCCTTCATTAAGGCGCTGGGTATGGCAGC 111
DB 491 TAACTTTCTAGTCCAAAGTGAAGTGAACCGCGCTTTATTAAGCAAGGCTGCACCCAGC 550
QY 112 CAGCATGTCTCACTGCCCGCCGAGACACAAACCCAGCATTTGAACACTGCACACGCGC 171

```

```

DB 551 AAGCATTCGTGACCTGCCACGCGGACCTGAGATTGGACACTTATCA-----GCCACTGC 605
QY 172 CATTCGCCAGAGAGCTGTGACCACCACTTC 202
DB 606 CAGTCCGCCAGAGAGCTGTGACCACCTTCC 636

RESULT 12
LOCUS      BC078889      1510 bp      mRNA      linear      ROD 03-AUG-2004
DEFINITION Rattus norvegicus insulin-like growth factor binding protein 1,
ACCESSION  BC078889
VERSION     BC078889.1 GI:50927646
KEYWORDS
SOURCE
ORGANISM   Rattus norvegicus (Norway rat)
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1510)
            Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stabileton,M., Soares,M.B., Bonaldi,M.J., Udell,T.B., Tsahuridu,L.,
            Scheetz,T.E., Brownstein,M.J., Ueffing,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Miallby,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Halyk,S.W.,
            Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fanhey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Scherchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL    12477932
REFERENCE  2 (bases 1 to 1510)
AUTHORS    Director MGC Project.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.ncl.nih.gov
            Contact: MGC help desk
            Email: cgaabp@mail.nih.gov
            Tissue Procurement: Howard Jacobs
            cDNA Library Preparation: Express Genomics
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) mcdpaxil@stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov
            Series: IRAX Project: 184 Row: e Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 6581079.
FEATURES   Location/Qualifiers
            source          1..1510
                        /organism="Rattus norvegicus"
                        /mol_type="mRNA"
                        /db_xref="taxon:10116"

```

```

gene
/clone="MGC:93595 IMAGE:7129185"
/tissue_type="Kidney", rac (Brown Norway)"
/clone_id="NIH_MGC_236"
/lab_host="DH10B"
/notes="Vector: pExpress1"
1..1510
/gene="Igfbp1"
/notes="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="LOCUSID:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
142..960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH7889.1"
/db_xref="GI:50927647"
/db_xref="LOCUSID:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
/translation="MPEPLTVSWPFLILISQVAVVAGAPQPMHCARCTARLELCP
PVPASCEPISRPAGCGCCPTCALPLGACGVATARCAQGLSCRALPGSRPLHALTRG
OGACVLEPAAPATISLSGSHBEAKAAVEDELASPEMTEBOLDSDHMAAPERD
OPIIMNAISTVSMRAREITDLKWKPECORELYVLRLLAAQCKADDELYKFLPN
CKNGFYHSKQETSIDGBAGLCWCYPMGSKKIPGSLERDGPNCQYFNVQN"

ORIGIN

Query Match      29.9%; Score 65.4; DB 10; Length 1510;
Best Local Similarity 98.5%; Pred. No. 7.9e-09;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 148 CGAGCATTAACACACGACGCGCATCTGCCAGAGACGTGTGACCACTTCGCGCTA 207
    |||||
Db 1 CGAGCATTAACACGACGCGCATCTGCCAGAGACGTGTGACCACTTCGCGCTA 60
    |||||

QY 208 CTAGCTA 214
    |||||
Db 61 CTATCTA 67

RESULT 13
AY095345      3886 bp      DNA      linear      PRI 30-DEC-2002
LOCUS      AY095345
DEFINITION      Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
ACCESSION      AY095345
VERSION      AY095345.1 GI:20853764
KEYWORDS
SOURCE
ORGANISM      Papio anubis (olive baboon)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE
AUTHORS      Kim,J.J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
TITLE      Regulation of insulin-like growth factor binding protein-1 promoter
activity by FMR and HOXA10 in primate endometrial cells
JOURNAL      Biol. Reprod. 68 (1), 24-30 (2003)
PUBMED      12493691
REFERENCE
AUTHORS      Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
TITLE      Direct Submission
JOURNAL      Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

FEATURES
source      1..3886
            Location/Qualifiers
            /organism="Papio anubis"
            /mol_type="genomic DNA"
            /db_xref="taxon:9555"
            <3655..>3886
            /gene="IGFBP-1"
gene

```

```

mRNA
<3655..>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655..>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/db_xref="GI:20853765"
/translation="MSEVPVAVRWLVLLLTVOYGVTVASAPWQCAPGSAEKALCPV
PASCSEVTRSGCGCCPMCALPLGACGVATAR"

CDS

ORIGIN

Query Match      28.6%; Score 62.6; DB 9; Length 3886;
Best Local Similarity 64.0%; Pred. No. 5.7e-08;
Matches 135; Conservative 0; Mismatches 59; Indels 17; Gaps 2;

QY 2 CACAAGCAAAACAACTATTATTGAACACGGGAGATCTAGACCGCTGCTGACATCAT 61
    |||||
Db 3366 CACTAGCAAAACAACTATTATTGAACACCTAGCTCTAGGCTGCCGCTGCCATCAT 3425
    |||||

QY 62 TAACCC-----GTGCTGCCGAGCGACCCCTTCATTAAGGCTGGGTAGGCCAGC 111
    |||||
Db 3426 TAACCTCCTGGTGAAGTGGCGCGGCTGTGCTTATTAAGGCGCGCTGTGTCCAGC 3485
    |||||

QY 112 CAGCATGTCACATGCGCCCGCGGACACCAAAACCAGGAGATTGACACTGCACAGGC 171
    |||||
Db 3486 AAGCATGGCGCCACCGCATCCATCCAGCAAGC-----ATTCGCGCGCGCGCGCGC 3538
    |||||

QY 172 CATCTGCCAGAGAGCTGTGACCACTTC 202
    |||||
Db 3539 CACCTCCAGAGACACTGCGCCACCGCTCC 3569
    |||||

RESULT 14
G67139      448 bp      DNA      linear      STS 18-SEP-2000
LOCUS      IGFBP1 X1.1 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION      G67139
ACCESSION      G67139.1 GI:10186730
VERSION      G67139.1
KEYWORDS      STS.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Cox,D.G., Bolland,C. and Canzian,F.
TITLE      Genome Survey
JOURNAL      Unpublished (2000)
COMMENT
Contact: Federico Canzian
Genome Analysis Group
International Agency for Research on Cancer
150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France
Tel: +33-4-72738698
Fax: +33-4-72738388
Email: canzian@iarc.fr
Primer A: TGCACTAGCAAAACAACCTT
Primer B: GGGTAGCTCCGAGCAGC
STS size: 448
Protocol:
Template:      50      ng
Primer:      4      um each
dNTPs:      2      mM each
MgCl2:      1.5-2.5      mM
Tag:      0.05      units
Total Vol:      25      ul
Buffer:
MgCl2:      1.5-2.5      mM
KCl:      50      mM
Tris-HCl:      20      mM
pH:      8.4.

```

FEATURES
source 1. 448
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind 1. .448
primer_bind 1. .20
complement(431. .448)
ORIGIN

Query Match 27.8%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 1.9e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAGCAAAACAACTTATTTTGAACACGGGATCTAGACGCTGCTGACAAATCAT 61
DB 3 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGTGGCGGCTGCAATCAT 62
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTATAGGCCCTGGGTATGCGCAGC 111
DB 63 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTTTATAGTGGCGCTGTGTCACAC 122
QY 112 CAGCATGTCTCACTGCCCGCCGAGACAAACCCAGCGACATTGAACACTGC-ACACGG 170
DB 123 GAGCATCGGCGCACCGCCATCC-----CATTCAGCGAGCATCTGCCCGCGCGCGCG 174
QY 171 CCATCTGCCCGAGAGAGCTGTGACCACTTC 202
DB 175 CCACCTCCAGAGAGCACTGGCGCACCGCTCC 206

RESULT 15
LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2394 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:21442452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Alvarez, C., Horne, D., Pera-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-ABR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1. 6128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN
Query Match 27.8%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 2e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAGCAAAACAACTTATTTTGAACACGGGATCTGCAAGCTGCGCCCTGACAAATCAT 61
DB 473 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGTGGCGGCTGCAATCAT 532
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTATAGGCCCTGGGTATGCGCAGC 111
DB 533 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTTTATAGTGGCGCTGTGTCACAC 592
QY 112 CAGCATGTCTCACTGCCCGCCGAGACAAACCCAGCGACATTGAACACTGC-ACACGG 170
DB 593 GAGCATCGGCGCACCGCCATCC-----CATTCAGCGAGCATCTGCCCGCGCGCGCG 644
QY 171 CCATCTGCCCGAGAGAGCTGTGACCACTTC 202

DB 645 CCACCTCCAGAGAGCACTGGCGCACCGCTCC 676

Search completed: September 1, 2005, 03:26:45
Job time : 1371.59 secs

This Page Blank (uspto)

Db	9	TGGGAGACCGACGATGTCATCTGCCCCGGGAGACACACCCAGCGACATTGAACACTG	68
Oy	164	CACACGGCCATCTGCCAGAGAGCTGTGACCAACCACTTCGGTACTAGCTA	214
Db	69	CACACGGCCGTCGTGCCAGAGAGCTGTGACCAACCACTTCGGTACTATATTA	119
RESULT 2			
LOCUS	BY783538		
DEFINITION	BY783538 RIKEN full-length enriched, 17.5 days embryo whole body		
ACCESSION	Mus musculus cDNA clone U930176D05 5', mRNA sequence.		
VERSION	BY783538		
KEYWORDS	EST.		
SOURCE	BY783538.1 GI:39710177		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 310)		
	Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanishi, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidini, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, U. and Hayashizaki, Y.		
TITLE	Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shohiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES			
SOURCE	Location/Qualifiers 1..310 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="U930176D05" /tissue_type="whole body" /dev_stage="17.5 days embryo" /clone_id="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	40.1%;	Score 87.8;	DB 6;
Best Local Similarity	92.9%;	Pred. No. 1.3e-15;	Length 310;
Matches	92;	Conservative 0;	Mismatches 7;
		Indels 0;	Gaps 0;
Oy	116	ATGTCACCTCTCCCGCGAGACAAACCCAGCGAGCTTGAACCTGACACGGGCATTC	175
Db	2	ATGTCACCTCTCCCGCGAGACAAACCCAGCGAGCTTGAACCTGACACGGGCATTC	61
Oy	176	TGCCAGAGACTGTACACCACTTCGCTACTAGCTA	214

DB	62	TGCCAGAGAGCTGTGACCACTATTCCTCA	100
RESULT 3			
LOCUS	BY794229		
DEFINITION	BY794229 RIKEN full-length enriched, 17.5 days embryo whole body		
ACCESSION	BY794229		
VERSION	BY794229.1		
KEYWORDS	GI:39720868		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 331)		
AUTHORS	Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Alizawa, K., Aizawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanabe, A., Hirozane-Kitahara, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kituch, N., Yoshiki, A., Kusabe, M., Gustincich, S., Beisel, K., Pavan, W., Aldinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakachi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Campar, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.		
TITLE	Targeting a complex transciptome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shinto-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES	Location/Qualifiers		
source	1..331 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="L930292H24" /tissue_type="whole body" /dev_stage="17.5 days embryo" /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	40.1%;	Score 87.8;	DB 6; Length 331;
Best Local Similarity	92.9%;	Pred. No. 1.3e-15;	
Matches	92;	Conservative 0;	Mismatches 7; Indels 0; Gaps 0;
QY	116	ATGATCCATGCGCCGCGAGACACAAACCCAGCAGCATTTAAACACTGCACACGGCCATC	175
DB	2	ATGATCCATGCGCCGCGAGACACACACCCAGCAGCATTTGAAACACTGCACACGGCCGTC	61
QY	176	TGCCAGAGAGCTGTGACCACTTCCGCTACATAGCTA	214
DB	62	TGCCAGAGAGCTGTGACCACTTGCACACTTGCACACTATCTCA	100
RESULT 4			
	BY771317		

LOCUS BY771317 354 bp mRNA linear EST 23-MAR-2004
 DEFINITION BY771317 RIKEN full-length enriched, 17.5 days embryo whole body
 ACCESSION Mus musculus cDNA clone U930052G15 5', mRNA sequence.
 VERSION BT771317
 KEYWORDS EST.
 SOURCE GI:39697955
 ORGANISM Mus musculus (house mouse)
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS 1 (bases 1 to 354)
 Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Suganara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Wataniki, A., Hirozane-Kitahara, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustinich, S., Beisel, K., Pavan, W., Aldini, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Pagliolini, M., Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
 TITLE Targeting a complex transcritpome: the construction of the mouse full-length cDNA encyclopedia
 JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)
 MEDLINE 22703353
 PUBMED 12819125
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome-gsc.riken.jp/
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome-gsc.riken.jp/) for further details.
 FEATURES
 source
 1..354
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="L930052G15"
 /tissue_type="whole body"
 /dev_stage="17.5 days embryo"
 /clone_1fb="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN
 Query Match 40.1%; Score 87.8; DB 6; Length 354;
 Best Local Similarity 92.9%; Pred. No. 1.4e-15;
 Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 116 ATGATTCACATGCCCCGCGAGACACAACCCAGGAGACATTGAACATCTGCACACGGCCATCT 175
 |||||
 DB 2 ATGGTCACTGCCCCGCGAGACACACACCCAGGAGACATTGAACATCTGCACACGGCCGTC 61
 |||||
 QY 176 TGCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 214
 |||||
 DB 62 TGCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 100
 |||||

RESULT 5
 LOCUS AT785818 401 bp mRNA linear EST 02-JUL-1999
 DEFINITION u17h05.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1888569 5' similar to gb:K15179 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION AT785818

VERSION AT785818.1 GI:5333534
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS 1 (bases 1 to 401)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: u17h05.x1
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972893
 Seg primer: custom primer used
 High quality sequence stop: 126.
 FEATURES
 source
 1..401
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1888569"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1fb="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGGGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TTCGAGCGCTTACG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGT, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTTAAGAAGTGGC and 3' end primer CGACCTGCAGCTCGACACA."
 primer CGACCTGCAGCTCGACACA."

ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 401;
 Best Local Similarity 92.9%; Pred. No. 2.8e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGATCACTGCCCCGCGAGACACAACCCAGGAGACATTGAACATCTGCACACGGCCATCT 176
 |||||
 DB 1 TGATCACTGCCCCGCGAGACACACACCCAGGAGACATTGAACATCTGCACACGGCCGCT 60
 |||||
 QY 177 GCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 214
 |||||
 DB 61 GCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 98
 |||||

RESULT 6
 LOCUS AT196314 480 bp mRNA linear EST 14-OCT-1998
 DEFINITION u17a07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1887828 5' similar to gb:M59316 rat1 INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION AT196314

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 480)
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The Mashu-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Mashu-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LIND, contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972152
 FEATURES
 source High quality sequence stop: 375.
 location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI0B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTCG); Site 2: DraIII (CACCATGTCG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTCG, 3' site CACCATGTCG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACCTCGACGACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 2.9e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGTGTCACTGCCCCCGGAGACACAAACCCAGGAGCATTTGAACATGCAACAGGCGCATCT 176
 DB 1 TGTGTCACTGCCCCCGGAGACACACACACCCAGGAGCATTTGAACATGCAACAGGCGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION u189f09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:W531e_rnal INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146 GI:4444281
 VERSION A1530146.1
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 706)
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The Mashu-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/Mashu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LIND, contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 FEATURES
 source High quality sequence stop: 479.
 location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI0B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTCG); Site 2: DraIII (CACCATGTCG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTCG, 3' site CACCATGTCG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACCTCGACGACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 3.1e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGTGTCACTGCCCCCGGAGACACAAACCCAGGAGCATTTGAACATGCAACAGGCGCATCT 176
 DB 1 TGTGTCACTGCCCCCGGAGACACACACACCCAGGAGCATTTGAACATGCAACAGGCGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION u21e1e07.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579 M. musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
 EST.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930344
 Seq primer: custom primer used
 High quality sequence stop: 395.
 FEATURES
 source
 1..765
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1481988"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCAGTGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCTACTGG), digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTGCAGCACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 765;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGGTCACTGCCCGCCGAGACACAAACCAGAGCATTTGAACACTGCACAGCGGCATCT 176
 DB 1 TGGTCACTGCCCGCCGAGACACACACCAGAGCATTTGAACACTGCACAGCGGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 9
 A1530313 785 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530313
 DEFINITION u191f01.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1889785 5' similar to gb:M59316.1 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
 VERSION A1530313.1 GI:4444448
 EST.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence stop: 459.
 FEATURES
 source
 1..785
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889785"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCAGTGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTGCAGCACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 785;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGGTCACTGCCCGCCGAGACACAAACCAGAGCATTTGAACACTGCACAGCGGCATCT 176
 DB 1 TGGTCACTGCCCGCCGAGACACACACCAGAGCATTTGAACACTGCACAGCGGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 10
 A1529939 799 bp mRNA linear EST 18-MAR-1999
 LOCUS A1529939
 DEFINITION u167c09.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1889392 5' similar to gb:M59316.1 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI529939 GI:4444074
 VERSION AI529939.1
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 799)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 FEATURES
 source location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pMB18-FL3; Site 1: DraIII (CACTGTCG); Site 2: DraIII (CACTGTCG); 1st strand cDNA was primed with an oligo(dt) primer [ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [CTGTGCTTACTG], digested and cloned into distinct DraIII sites of the pMB18-FL3 vector (5' site CACTGTCG, 3' site CACTGTCG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTCTAAAGCTGCG and 3' end primer CGACTGCACTGACGACACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCCACTGCCCCCGGAGACACAAACCCAGGAGCATTTGAACATGCACAGGCCCATCT 176
 Db 1 TGGTCCACTGCCCCGGGAGACACACACCCAGGAGATTGAACACTGACACAGGCCGCTCT 60
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 11
 CA478518 811 bp mRNA linear EST 09-MAR-2004
 LOCUS CA478518
 DEFINITION AGENCOURT_10789306 NIH_MGC_152 Mus musculus cDNA clone
 IMAGE:5766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE NIH-MGC Catalog
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LMK00138 row: m column: 23
 High quality sequence stop: 536.
 FEATURES
 source location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCTCGAGGCGCGCCAGCACACATTTGACGAAGAAGCGGCTTTTCTTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen. Library amplification 16 cycles. Library constructed by Mark Bittinger in the Bradfield Laboratory (McCardle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCCACTGCCCCCGGAGACACAAACCCAGGAGCATTTGAACATGCACAGGCCCATCT 176
 Db 39 TGGTCCACTGCCCCGGGAGACACACACCCAGGAGATTGAACACTGACACAGGCCGCTCT 98
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 136
 RESULT 12
 A1790802 848 bp mRNA linear EST 02-JUL-1999
 LOCUS A1790802
 DEFINITION uk28b10.y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 IMAGE:1970299 5', similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 848)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28p10.x1
Contact: Marra M/Waehu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039
Seq primer: custom primer used
High quality sequence stop: 514.

FEATURES

source

1.848
Location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mKia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CAGCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TTTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CAGCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGCTCTTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 39.6%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 3.2e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGTGTCACCTGCCCCCGAGACACAAACCGAGGACATTGAACACTGACACAGGCACT 176
|||||
DB 1 TGTTCACCTGCCCCCGAGACACACACACCGAGGACATTGAACACTGACACAGGCGTCT 60
|||||
QY 177 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 214
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 98
|||||

RESULT 13
A1528304 852 bp mRNA linear EST 18-MAR-1999
LOCUS A1528304
DEFINITION u195g10.y1 Sugano mouse liver m1a Mus musculus cDNA clone
IMAGE:1890210.5 similar to gb:W5316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1528304
VERSION A1528304.1 GI:4442439
KEYWORDS EST.
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 852)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Paton,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

JOURNAL

Unpublished (1999)

Contact: Marra M/Waehu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534
Seq primer: custom primer used
High quality sequence stop: 478.

FEATURES

source

1.852
Location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver m1a"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CAGCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TTTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CAGCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGCTCTTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 39.6%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 3.2e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGTGTCACCTGCCCCCGAGACACAAACCGAGGACATTGAACACTGACACAGGCACT 176
|||||
DB 1 TGTTCACCTGCCCCCGAGACACACACACCGAGGACATTGAACACTGACACAGGCGTCT 60
|||||
QY 177 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 214
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 98
|||||

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998
LOCUS A1196154
DEFINITION u193d08.y1 Sugano mouse liver m1a Mus musculus cDNA clone
IMAGE:1887665.5 similar to gb:W5316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1196154
VERSION A1196154.1 GI:3748760
KEYWORDS EST.
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 605)
Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMNI Mouse EST Project

JOURNAL COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:971987

Seq primer: custom primer used
 High quality sequence stop: 420.

FEATURES

source

1..605
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887663"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end primer CGACTGTGACCTGACGACA."

ORIGIN

Query Match

Best Local Similarity 38.9%; Score 85.2; DB 1; Length 605;
 Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 117 TGGTCCACTGCCCGCGAGACACAAACCAGAGGATTAACACTGCACAGCGCATCT 176
 |||||
 DB 1 TGGTCCACTGCCCGCGAGACACACCCAGAGATTGAACACTGCACAGCGTGTCT 60

QY 177 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
 |||||
 DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 15

A1785039

LOCUS ui73a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 DEFINITION IMAGE:1888018.5' similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION

A1785039

A1785039.1 GI:5332755

EST.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished (1999)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:971987
 Seq primer: custom primer used
 High quality sequence stop: 420.
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887663"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end primer CGACTGTGACCTGACGACA."

COMMENT

Other ESTs: ui73a06.x1
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972342
 Seq primer: custom primer used
 High quality sequence stop: 353.

FEATURES

source

1..380
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1888018"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end primer CGACTGTGACCTGACGACA."

ORIGIN

Query Match

Best Local Similarity 37.5%; Score 82.2; DB 1; Length 380;
 Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 120 TCCACTGCCCGCGAGACACAAACCAGAGGATTAACACTGCACAGCGCATCTGCC 179
 |||||
 DB 1 TCCACTGCCCGCGAGAGACACACCCAGAGATTGAACACTGCATACAGCGCGTGTGCC 60

QY 180 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
 |||||
 DB 61 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 95

Search completed: September 1, 2005, 06:56:46
 Job time : 1673.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 ; Search time 637.824 Seconds
(without alignment)
2248.600 Million cell updates/sec

Title: US-09-972-916B-2

Perfect score: 219
Sequence: 1 tcacagcgaacaaacta.....ttccgctactagctacgcgc 219

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219	100.0	219	US-09-972-916A-2	Sequence 2, Appli
2	219	100.0	270	US-09-972-916A-3	Sequence 3, Appli
3	219	100.0	321	US-09-972-916A-4	Sequence 4, Appli
4	219	100.0	423	US-09-972-916A-5	Sequence 5, Appli
5	219	100.0	472	US-09-972-916A-6	Sequence 6, Appli
6	76.4	34.9	1500	US-09-917-800A-1608	Sequence 1608, Ap
7	76.4	34.9	1500	US-10-388-934-5	Sequence 5, Appli

8	76.4	34.9	1500	17	US-10-191-803-73	Sequence 73, Appli
9	76.4	34.9	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
10	60.8	27.8	6128	9	US-09-880-107-2393	Sequence 2393, Ap
11	60.8	27.8	6128	22	US-10-756-149-1484	Sequence 1484, Ap
12	60.8	27.8	9173	22	US-10-893-315-126	Sequence 126, App
13	60.8	27.8	9174	22	US-10-893-315-160	Sequence 160, App
14	37.8	17.3	7061	15	US-10-311-455-970	Sequence 970, App
15	37.8	17.3	7061	17	US-10-221-613-148	Sequence 148, App
16	35.2	16.1	1687	20	US-10-739-930-5461	Sequence 5461, App
17	35	16.0	769	13	US-10-027-532-164336	Sequence 164336, Sequence 5746, Appl
18	35	16.0	769	17	US-10-027-532-164336	Sequence 164336, Sequence 5746, Appl
19	35	16.0	3895	14	US-10-011-585A-76	Sequence 76, Appl
20	34	15.5	49979	19	US-10-741-600-17995	Sequence 17995, A
21	34	15.5	49979	21	US-10-437-863-79182	Sequence 79182, A
22	32.4	14.8	1389	19	US-10-437-863-79182	Sequence 79185, A
23	32.4	14.8	1584	19	US-10-322-281-492	Sequence 492, App
24	32	14.6	76180	19	US-10-322-281-492	Sequence 31, Appl
25	31.6	14.4	1129	20	US-10-343-803-31	Sequence 2017, Ap
26	31.6	14.4	1288	15	US-10-017-161-2017	Sequence 1663, Ap
27	31.6	14.4	1288	17	US-10-292-798-1663	Sequence 1663, Ap
28	31.6	14.4	1348	17	US-10-120-988-404	Sequence 42, App
29	31.6	14.4	1513	18	US-10-072-012-29	Sequence 169, App
30	31.6	14.4	117750	19	US-10-367-094-169	Sequence 169, App
31	31.4	14.3	360	20	US-10-425-115-28332	Sequence 28332, A
32	31.4	14.3	1134	17	US-10-282-122A-33251	Sequence 33251, A
33	31.2	14.2	668	20	US-10-363-345A-40493	Sequence 40493, A
34	31.2	14.2	668	20	US-10-363-345A-40494	Sequence 40494, A
35	31.2	14.2	668	21	US-10-363-483A-40493	Sequence 40494, A
36	31.2	14.2	668	21	US-10-363-483A-40494	Sequence 2769, Ap
37	31.2	14.2	955	20	US-10-653-047-2769	Sequence 1314, Ap
38	31.2	14.2	2019	15	US-10-156-761-1314	Sequence 24, App
39	31.2	14.2	6226	17	US-10-221-613-254	Sequence 1, Appl
40	31.2	14.2	9025608	15	US-10-156-761-1	Sequence 8255, A
41	31	14.2	466	22	US-10-972-079-82551	Sequence 51, Appl
42	31	14.2	2190	21	US-10-696-909A-51	Sequence 34, Appl
43	31	14.2	2247	16	US-10-119-428-31	Sequence 31, Appl
44	31	14.2	2247	17	US-10-291-172-31	Sequence 31, Appl
45	31	14.2	2247	18	US-10-221-278-31	

ALIGNMENTS

RESULT 1
US-09-972-916A-2
; Sequence 2, Application US/09972916A
; Patent No. US20020107198A1
GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
US-09-972-916A-2

Query Match 100.0%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 1e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGCAAACTTATTTGAACACGGGATCTTACAGCGCTGCGCATCA 60
DB 1 TCACAGCAAACTTATTTGAACACGGGATCTTACAGCGCTGCGCATCA 60
QY 61 TTACCGGTGTCGCGACGCGCTTATTAAGCGCGGTATGCGCAGCAGCATGT 120

Db 61 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 120
QY 121 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 180
Db 121 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 180
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219

RESULT 2
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 100.0%; Score 219; DB 9; Length 270;
Best Local Similarity 100.0%; Pred. No. 1,1e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 60
Db 52 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 111
QY 61 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 120
Db 112 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 171
QY 121 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 180
Db 172 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 231
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 232 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 270

RESULT 3
US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 100.0%; Score 219; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 1,1e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 60
Db 103 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 162
QY 61 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 120
Db 163 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 222
QY 121 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 180
Db 223 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 282
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 283 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 321

RESULT 4
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 100.0%; Score 219; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 1,2e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 60
Db 154 TCACAGCAAAACAACTTATTTTGAACACGCGGATCCTAGACGCTGCGCTGACATCA 213
QY 61 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 120
Db 214 TTAACCGTGTGCTCCGAGACCGCTTCAATGAAGCCCTGGGTATGCGCAGCATGCT 273
QY 121 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 180
Db 274 CCACCTGCGCGCGAGACACAACCCGAGCATTTGAACCTGCAACAGCGCATCTGCGCC 333
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 334 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 372

RESULT 5
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A

```
/ CURRENT FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 6
/ LENGTH: 423
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match          100.0%; Score 219; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No.1.2e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAAACAACAACTTATTTTGAACGCGGATCTAGACGCTGCTTACATCA 60
DB 205 TCACAAACAACAACTTATTTTGAACGCGGATCTAGACGCTGCTTACATCA 264
QY 61 TTAACCGCTGCTCCGAGCCAGCCCTTCATTAAGGCGCTGATGCGCCAGCAGATGAT 120
DB 265 TTAACCGCTGCTCCGAGCCAGCCCTTCATTAAGGCGCTGATGCGCCAGCAGATGAT 324
QY 121 CCACTGCGCGCCGAGACAAACCCAGCAGCATTTGAACGCGCCATCTGCCCC 180
DB 325 CCACTGCGCGCCGAGACAAACCCAGCAGCATTTGAACGCGCCATCTGCCCC 384
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGAGCCGC 219
DB 385 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGAGCCGC 423

RESULT 6
US-09-917-800A-1608
/ Sequence 1608, Application US/09917800A
/ Patent No. US20020119462A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Porter, Mark
/ APPLICANT: Johnson, Kory
/ APPLICANT: Castle, Arthur
/ APPLICANT: Elashoff, Michael
/ APPLICANT: Gene Logic, Inc.
/ TITLE OF INVENTION: Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5038-US
/ CURRENT APPLICATION NUMBER: US/09/917,800A
/ CURRENT FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,040
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,880
/ PRIOR FILING DATE: 2000-11-02
/ PRIOR APPLICATION NUMBER: US 60/290,029
/ PRIOR FILING DATE: 2001-05-11
/ PRIOR APPLICATION NUMBER: US 60/290,645
/ PRIOR FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: US 60/292,336
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: US 60/295,798
/ PRIOR FILING DATE: 2001-06-06
/ PRIOR APPLICATION NUMBER: US 60/297,457
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,884
/ PRIOR FILING DATE: 2001-06-19
/ PRIOR APPLICATION NUMBER: US 60/303,459
/ PRIOR FILING DATE: 2001-07-09
/ NUMBER OF SEQ ID NOS: 1740
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 1608
/ LENGTH: 1500
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
```

```
/ OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match          34.9%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No.6.2e-17;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTCCGCTACTAGCTA 214
DB 68 CACTCCGCTACTACTA 85

RESULT 7
US-10-388-934-5
/ Sequence 5, Application US/10388934
/ Publication No. US20040005547A1
/ GENERAL INFORMATION:
/ APPLICANT: Boess, Franziska
/ APPLICANT: Suter-Dick, Laura
/ APPLICANT: Wolf, Delfie
/ TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
/ FILE REFERENCE: 21199
/ CURRENT APPLICATION NUMBER: US/10/388,934
/ CURRENT FILING DATE: 2003-03-14
/ PRIOR APPLICATION NUMBER: 02005336.9
/ PRIOR FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 02015657.6
/ PRIOR FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 862
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 5
/ LENGTH: 1500
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-5

Query Match          34.9%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No.6.2e-17;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTCCGCTACTAGCTA 214
DB 68 CACTCCGCTACTACTA 85

RESULT 8
US-10-191-803-73
/ Sequence 73, Application US/10191803
/ Publication No. US20040014040A1
/ GENERAL INFORMATION:
/ APPLICANT: MENDRICK, Donna
/ APPLICANT: PORTER, Mark
/ APPLICANT: JOHNSON, Kory
/ APPLICANT: HIGGS, Brandon
/ APPLICANT: CASTLE, Arthur
/ APPLICANT: ELASHOFF, Michael
/ TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5090US
/ CURRENT APPLICATION NUMBER: US/10/191,803
/ CURRENT FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: US 60/303,819
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/305,623
/ PRIOR FILING DATE: 2001-07-17
/ PRIOR APPLICATION NUMBER: US 60/369,351
```

? PRIOR FILING DATE: 2002-04-03
 ? PRIOR APPLICATION NUMBER: US 60/377,611
 ? PRIOR FILING DATE: 2002-05-06
 ? NUMBER OF SEQ ID NOS: 1140
 ? SOFTWARE: PatentIn Ver. 2.1
 ? SEQ ID NO 73
 ? LENGTH: 1500
 ? TYPE: DNA
 ? ORGANISM: *Rattus norvegicus*
 ? FEATURE:
 ? OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_013144
 ? US-10-191-803-73

Query Match	34.9%;	Score 76.4;	DB 17,	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 6.2e-17;		
Matches 77; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Dy 137 CACAAACCAGCGGAGCATTTGAACACTGCACACGGCCATCTGTGCCAGAGAGCTGTGACCAAC 196
Db 8 CACAAACCAGCGGAGCATTTGAACACTGCACACGGCCATCTGTGCCAGAGAGCTGTGACCAAC 67

Qy	197	CACTTCCGCTACTAGCTA	214
Db	68	CACTTCCGCTACTATCTA	85

RESULT 9
US-10-152-319A-1613

```

; Sequence 1613, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:

```

```
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
```

```

; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael

```

```

; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A

```

```

; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523

```

;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/298,925
;; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/303,810
 ;
 ; PRIOR FILING DATE: 2001-07-10
 ;
 ; PRIOR APPLICATION NUMBER: US 60/303,807
 ;

; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,808
 ; PRIOR FILING DATE: 2001-07-10

;; PRIOR APPLICATION NUMBER: US 60/315,047
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: US 60/324,928

; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US 60/330,867
 ; PRIOR FILING DATE: 2001-11-01

; PRIORITY APPLICATION NUMBER: US 60/330,462
 ; PRIORITY FILING DATE: 2001-10-22
 ; Remaining Prior Application data removed - See File Wrapper or PAM

```

; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1613

```

```

;
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus

```

```

;
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_013144
; US-10-152-319A-1613

```

Query Match 34.9%; Score 76.4; DB 18; Length 1500;

Best Local Similarity 98.7%; Pred. No. 6, 2e-17;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Oy 137 CACAAACCCAGCGCATTTGAACA CTGCA CGGCATCTGCCAGAGAGCTGTGACCAC 196
| | | | |
D8 8 CACAAACCCAGCGCATTTGAACA CTGCA CGGCATCTGCCAGAGAGCTGTGACCAC 67
| | | | |

Qy	197	CACTCCGCTACTAGCTA	214
Db	68	CACTCCGCTACTATCTA	85

RESULT 10
US-09-880-107-2393

```

; Sequence 2393, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe

```

```

; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO

```

; CURRENT APPLICATION NUMBER: US/09/880,107
 ;
 ; CURRENT FILING DATE: 2001-06-14
 ;
 ; PRIOR APPLICATION NUMBER: US 60/211,379

;; PRIOR FILING DATE: 2000-06-14
;; PRIOR APPLICATION NUMBER: US 60/237,054
;; PRIOR FILING DATE: 2000-10-02

```

; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2393

```

```

; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M745877
US-09-880-107-2393

```

Query Match	27.8%	Score 60.8;	DB 9;	Length 6128;
Best Local Similarity	66.5%	Pred. No. 4e-11;		
Matches 141; Conservative	0;	Mismatches 52;	Indels 19;	Gaps 3.

2 CACAGCAAACAACTTATTTTGACACGGGATCTTAGCAGCGTGCCTGACATCAT 61

Db	473	CACTCAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGCGCTGCCAATCAT	5122
Oy	62	TAAACC-----GTGCTGCCGAGCAGCCCTTCATAGAAGCCCTGGTATGCGCAGC	111

Db
533 TAACTCTGTCAGATGGCGCGGCGCTGTGCCCTTTATTAAGGTGGCGCGCTGTCCAGC 592

QY 112 CAGCATGTCCTACCTGCCCGCGAGACACAAACCGAGCATTTGAACATCTC-ACA CGG 170

Db 593 GAGCATCGGCACCGGCATCTC-----CATCAGCGAGCATCTGCGCGCGCGCGCG 644

171 CCATCTGCCAGAGGCTGTGACCAACCACTTC 202

Db 645 CCACCCTCCAGAGCACTGGCCACCGCTCC 676

RESULT 11
US-10-756-149-1484
: Sequence 1484, Application US/10756149

Publication No. US20050181375A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha

APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODIFIATORS OF METASTATIC CANCER

```

; FILE REFERENCE: file
;
; CURRENT APPLICATION NUMBER: US/10/756,149
;
; CURRENT FILING DATE: 2004-01-12

```

NUMBER OF SEQ ID NOS: 5818

SOFTWARE: Patentin version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-1484

Query Match 27.8%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAAGCAAAACAACTATTATTGAAACAGCGGATCTGACAGCTGCGCTGACATCAT 61
DB 473 CACTAGCAAAACAACTATTATTGAAACACTGACCTCTGACCGCGCGCTGCAATCAT 532
QY 62 TAACCC-----GTGCTGCCAGAGCCGCTTCATTAAGCCCTGGGTATGCCAGC 111
DB 533 TAACCTCTGCTGCAAGTGCGCGGCTGTGCTTATTAAGTGCGCGCTGTGTCCAGC 592
QY 112 CAGCATGTGCTCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
DB 593 GAGCATGCGCCACCGCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGC 644
QY 171 CCATCTGCCAGAGAGCTGTGACCACTTC 202
DB 645 CCACCTCCCAAGAGACACTGCGCCACGCTCC 676

RESULT 12
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-126

Query Match 27.8%; Score 60.8; DB 22; Length 9173;
Best Local Similarity 66.5%; Pred. No. 4.4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAAGCAAAACAACTATTATTGAAACAGCGGATCTGACAGCTGCGCTGACATCAT 61
DB 1877 CACTAGCAAAACAACTATTATTGAAACACTGACCTCTGACCGCGCTGCCAATCAT 1936
QY 62 TAACCC-----GTGCTGCCAGAGCCGCTTCATTAAGCCCTGGGTATGCCAGC 111
DB 1937 TAACCTCTGCTGCAAGTGCGCGGCTGTGCTTATTAAGTGCGCGCTGTGTCCAGC 1996
QY 112 CAGCATGTGCTCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
DB 1997 GAGCATGCGCCACCGCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGC 2048
QY 171 CCATCTGCCAGAGAGCTGTGACCACTTC 202
DB 2049 CCACCTCCCAAGAGACACTGCGCCACGCTCC 2080

RESULT 13
US-10-893-315-160

Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-160

Query Match 27.8%; Score 60.8; DB 22; Length 9174;
Best Local Similarity 66.5%; Pred. No. 4.4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAAGCAAAACAACTATTATTGAAACAGCGGATCTGACAGCTGCGCTGACATCAT 61
DB 1877 CACTAGCAAAACAACTATTATTGAAACACTGACCTCTGACCGCGCTGCCAATCAT 1936
QY 62 TAACCC-----GTGCTGCCAGAGCCGCTTCATTAAGCCCTGGGTATGCCAGC 111
DB 1937 TAACCTCTGCTGCAAGTGCGCGGCTGTGCTTATTAAGTGCGCGCTGTGTCCAGC 1996
QY 112 CAGCATGTGCTCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
DB 1997 GAGCATGCGCCACCGCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGC 2048
QY 171 CCATCTGCCAGAGAGCTGTGACCACTTC 202
DB 2049 CCACCTCCCAAGAGACACTGCGCCACGCTCC 2080

RESULT 14
US-10-311-455-970/C
; Sequence 970, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 970
; LENGTH: 7061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-311-455-970

Query Match 17.3%; Score 37.8; DB 15; Length 7061;
Best Local Similarity 73.8%; Pred. No. 0.0086;
Matches 48; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Oy	2	CACACGCAAAACAACTATTGTAACACGGGAACTTAGACAGCTGGCCCGACAATCAT	61
Db	2185	CACATACAAAACAACTATTTTTAAACACTCACTCTTAAGTACGACGCTACCAATCAT	2126
Oy	62	TAACC	66
Db	2125	TAACC	2121

RESULT 15
THE 10-231

```

US-10-221-613-148/c
; Sequence 148, Application US/10221613
; Publication No. US20040029123A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
FILE REFERENCE: 5013.1004
CURRENT APPLICATION NUMBER: US/10/221.613
CURRENT FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: PCT/EP01/02945
DE 10013847.00
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-03-15
;
2000-03-15
;
2000-04-06
2000-04-07
2000-06-30
2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 148
; LENGTH: 7061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-148

```

Query Match 17.3%; Score 37.8; DB 17; Length 7061;
Best Local Similarity 73.8%; Pred. No. 0.0086;
Matches 48; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Oy	2	CACAGCAAAA	CAAACTATT	TTGGAACG	GGGAACT	CAGACAG	CTGGCC	TGCACAT	AT	61
Db	2185	CACATACAAA	CAAACTATT	TTTAAACAC	TCAACTC	CTTACGT	AGACGC	TACCAAT	AT	2126
Oy	62	TAAAC	66							
Db	2125	TAAAC	2121							

Search completed: September 1, 2005, 16:44:10
Job time : 647.824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 69.1649 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916B-2

Perfect score: 219

Sequence: 1 tcacacagcaaacactta.....ttccgctactagctagccgc 219

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
6: /cgn2_6/prodata/1/ina/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.8	27.8	9173	4 US-09-949-001-30	Sequence 30, App1
2	60.8	27.8	9174	4 US-09-949-001-36	Sequence 36, App1
3	35	16.0	601	4 US-09-949-016-19194	Sequence 19194, A
4	35	16.0	601	4 US-09-949-016-73784	Sequence 73784, A
5	35	16.0	2393	4 US-09-949-016-2153	Sequence 2153, App1
6	35	16.0	2408	4 US-09-949-016-74	Sequence 74, App1
7	35	16.0	7561	4 US-09-949-016-11816	Sequence 11816, A
8	35	16.0	7562	4 US-09-949-016-13895	Sequence 13895, A
9	32.2	14.7	4447	2 US-08-304-309-3	Sequence 3, App1
10	32.2	14.7	4447	2 US-08-991-942-3	Sequence 3, App1
11	32	14.6	36180	4 US-09-949-016-11745	Sequence 11745, A
12	32	14.6	36181	4 US-09-949-016-16163	Sequence 16163, A
13	31.6	14.4	1348	4 US-09-949-016-404	Sequence 404, App
14	31.6	14.4	99748	4 US-09-949-016-11990	Sequence 11990, A
15	31.6	14.4	99749	4 US-09-949-016-16518	Sequence 16518, A
16	31.2	14.2	364	4 US-09-270-767-29006	Sequence 29006, A
17	31.2	14.2	601	4 US-09-949-016-168692	Sequence 168692, A
18	31.2	14.2	601	4 US-09-949-016-168693	Sequence 168693, A
19	31.2	14.2	761	4 US-09-270-767-13109	Sequence 13109, A
20	31	14.2	1888	4 US-09-949-016-1341	Sequence 1341, App
21	31	14.2	13145	4 US-09-949-016-11883	Sequence 11883, A
22	30.6	14.0	897	4 US-09-902-540-4899	Sequence 4899, App
23	30.6	14.0	28493	4 US-09-902-540-1241	Sequence 1241, App
24	30.2	13.8	40181	4 US-09-949-016-17016	Sequence 17016, A
25	29.8	13.6	1056	3 US-09-067-782A-1	Sequence 1, App1
26	29.4	13.4	488	4 US-09-621-976-2763	Sequence 2763, App
27	29.2	13.3	1101	4 US-09-205-258-119	Sequence 119, App

c	28	29	13.2	790	3	US-08-998-416-479	Sequence 479, App
c	29	29	13.2	2194	4	US-09-270-767-10411	Sequence 10411, A
c	30	29	13.2	2205	4	US-09-252-991A-11217	Sequence 11217, A
c	31	29	13.2	2454	4	US-09-252-991A-11038	Sequence 11038, A
c	32	29	13.2	13489	4	US-09-949-016-15911	Sequence 15911, A
c	33	29	13.2	20235	4	US-07-642-734C-3	Sequence 3, App1
c	34	29	13.2	20235	3	US-08-439-009A-3	Sequence 3, App1
c	35	28.8	13.2	439	3	US-09-067-782A-3	Sequence 3, App1
c	36	28.8	13.2	777	4	US-09-902-540-3367	Sequence 3367, App
c	37	28.8	13.2	1383	1	US-08-484-044-1	Sequence 1, App1
c	38	28.8	13.2	3182	1	US-08-484-044-11	Sequence 11, App1
c	39	28.8	13.2	3759	1	US-08-252-966B-11	Sequence 11, App1
c	40	28.8	13.2	3785	1	US-08-252-966B-17	Sequence 17, App1
c	41	28.8	13.2	11613	1	US-08-484-044-10	Sequence 10, App1
c	42	28.8	13.2	17125	4	US-09-902-540-1158	Sequence 1158, App
c	43	28.6	13.1	1140	4	US-09-252-991A-901	Sequence 901, App
c	44	28.6	13.1	1615	4	US-09-799-451-116	Sequence 116, App
c	45	28.6	13.1	1765	4	US-09-799-451-117	Sequence 117, App

ALIGNMENTS

RESULT 1	US-09-949-001-30	
/	Sequence 30, Application US/09949001	
/	Patent No. 6825336	
/	GENERAL INFORMATION:	
/	APPLICANT: VENTER, J. Craig et al.	
/	TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED	
/	FILE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF	
/	FILE REFERENCE: C1000789	
/	CURRENT APPLICATION NUMBER: US/09/949,001	
/	PRIOR FILING DATE: 2003-01-15	
/	PRIOR APPLICATION NUMBER: 60/231,323	
/	PRIOR FILING DATE: 2000-09-08	
/	NUMBER OF SEQ ID NOS: 848	
/	SOFTWARE: FastSeq for Windows Version 4.0	
/	SEQ ID NO 30	
/	LENGTH: 9173	
/	TYPE: DNA	
/	ORGANISM: Human	
/	US-09-949-001-30	
Query Match	27.8%; Score 60.8; DB 4; Length 9173;	
Best Local Similarity	66.5%; Pred. No. 8.8e-10;	
Matches 141; Conservative	0; Mismatches 52; Indels 19; Gaps 3;	
QY	2 CACAAAGCAAACTTTTGAACAGCGGATCTTACGACCGCTGCCTGACAAATCAT	61
DB	1877 CACTAGCAAACTTATTTGAACAGCTGCTGCGGCTGCCTGCAATCAT	1936
QY	62 TAACCC-----GTGCTGCCGAGCGACGCTTCAAGCCCTGGATATGCGCAGC	111
DB	1937 TAACCTCTGTCGACAGTGGCGCGCTGTGCTTTAAGTCCGCTGTCTCAGC	1996
QY	112 CAGCATGTCTCACTCCCGCGGAGACACAAACCCGAGCATTTGAACATTCG-ACAGCG	170
DB	1997 GAGCATCGGACCAACCCGATCC-----CATTCGAGCATCTGCGCGCGCGCGCG	2048
QY	171 CCATCTGCCAGAGCGTGTGACCAACCATTC	202
DB	2049 CCACCTCCAGAGGACGACGACCAACCATTC	2080
RESULT 2	US-09-949-001-36	
/	Sequence 36, Application US/09949001	
/	Patent No. 6825336	
/	GENERAL INFORMATION:	
/	APPLICANT: VENTER, J. Craig et al.	
/	TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED	
/	FILE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF	

FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match 27.8%; Score 60.8; DB 4; Length 9174;
Best Local Similarity 66.5%; Pred. No. 8.8e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAAACAAACAACTTATTTTGAACACGCGGATCTTACGACGCTGCTGACAACTAT 61
DB 1877 CACTAGCAAAACAACTTATTTTGAACACTGACCTCTGCGCGGCTGCAATCAT 1936
QY 62 TAACCC-----GTCTGCCGACGACGCTTCAATAGCCCTGGGTATGGCCAGC 111
DB 1937 TAACCTCTGTGCAAGTGGCGGCTGTGCTTATAGGTGGCGCTGTGTCCAGC 1996
QY 112 CAGCATGTCTCACTGCGCGCGGACACAAACCGACGACATTTGAACACTGC-ACAGG 170
DB 1997 GAGCATCGGCCACCGCATCC-----CATCCAGCGACATCTGCGCGCGCGCGG 2048
QY 171 CCATCTGCCGAGAGAGCTGTGACACCACTTC 202
DB 2049 CCACCTCCGAGAGACACTGGCCACGCTCC 2080

RESULT 3
US-09-949-016-19194/c
Sequence 19194, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19194
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-19194

Query Match 16.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 79 CCAGCCCTTCAATAAGCCCTGGGTATGGCCAGCCAGACATGCTCCACTGCGCCGACAGACA 138
DB 478 CAGGCTTCTCGGGGCGCTGTCTCTTCATCCAGCTTAGCCACTCCAGAGTTAGGCC 419
QY 139 CAAACCCGACGAGCATTTGAACACTGCACACGCGCATCTTCCGACGAGAGCTGTGACCA 198
DB 418 CCATTCAGAGACATCCACCGGATCCAGAGCTTGCACACCACTTCATGATGATGCC 359
QY 199 CTT 201
|||

DB 358 CTT 356

RESULT 4
US-09-949-016-73784/c
Sequence 73784, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73784
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-73784

Query Match 16.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 79 CCAGCCCTTCAATAAGCCCTGGGTATGGCCAGCCAGACATGCTCCACTGCGCCGACAGACA 138
DB 478 CAGGCTTCTCGGGGCGCTGTCTCTTCATCCAGCTTAGCCACTCCAGAGTTAGGCC 419
QY 139 CAAACCCGACGAGCATTTGAACACTGCACACGCGCATCTTCCGACGAGAGCTGTGACCA 198
DB 418 CCATTCAGAGACATCCACCGGATCCAGAGCTTGCACACCACTTCATGATGATGCC 359
QY 199 CTT 201
|||
DB 358 CTT 356

RESULT 5
US-09-949-016-2153/c
Sequence 2153, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2153
LENGTH: 2393
TYPE: DNA
ORGANISM: Human
US-09-949-016-2153

Query Match 16.0%; Score 35; DB 4; Length 2393;
Best Local Similarity 55.3%; Pred. No. 0.18;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;


```

; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..3162
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4447
; OTHER INFORMATION: /product= "pig DPD"
;
; US-08-304-309-3
;
; Query Match
; Best Local Similarity 49.7%; Score 32.2; DB 2; Length 4447;
; Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
;
; QY 45 GCTGCCCTGACATCTTAACCCGCTGCTGCCGAGCCGCTTCATTAAGCCCTGGGTAT 104
; DB 2275 GATGGTGTATACGACCAACACGCGTCTCAGCTCTCATGAGATTAAGCCGATGGCAG 2334
;
; QY 105 GGCAGCGCAGCATGTGCCCTGCCCGCCGAGACAAACCAGCAGCATTTGAACACTGC 164
; DB 2335 CCTGGCGACGGGTGGGTGCTGGCAAGCGGACTATACGAGTGTCTGGCAGCGCC 2394
;
; QY 165 ACACGGCCATCTGCCAGAGAGCTGTGACCACTTCCTGGCTACT 209
; DB 2395 ATCAGACCAATTCCTTTGAGAGCTGTGACCACTTCCTGGCT 2439
;
; RESULT 10
; US-08-991-942-3
; Sequence 3, Application US/08991942
; Patent No. 6015673
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROXYIMIDINE DEHYDROENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza

```

```

; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,942
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..3162
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4447
; OTHER INFORMATION: /product= "pig DPD"
;
; US-08-991-942-3
;
; Query Match
; Best Local Similarity 49.7%; Score 32.2; DB 3; Length 4447;
; Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
;
; QY 45 GCTGCCCTGACATCTTAACCCGCTGCTGCCGAGCCGCTTCATTAAGCCCTGGGTAT 104
; DB 2275 GATGGTGTATACGACCAACACGCGTCTCAGCTCTCATGAGATTAAGCCGATGGCAG 2334
;
; QY 105 GGCAGCGCAGCATGTGCCCTGCCCGCCGAGACAAACCAGCAGCATTTGAACACTGC 164
; DB 2335 CCTGGCGACGGGTGGGTGCTGGCAAGCGGACTATACGAGTGTCTGGCAGCGCC 2394
;
; QY 165 ACACGGCCATCTGCCAGAGAGCTGTGACCACTTCCTGGCTACT 209
; DB 2395 ATCAGACCAATTCCTTTGAGAGCTGTGACCACTTCCTGGCT 2439
;
; RESULT 11
; US-09-949-016-11745
; Sequence 11745, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

; Patent No. 6743619
;
; GENERAL INFORMATION:
;
; APPLICANT: Tang, Y. Tom

```

72 TGCAGACCAACCCTTCATTAAGGCCCTGAGTATGACCAGCCAGCATGTGCCATGCCCCC 131

Db 33666 TTCTGAGCAAAACACACACAGCCCAATTGTATACACACCAGAAAGTGGAATGATGGG 33607
 QY 132 CGAGACACAAACCCAGCCGAGCATTTGAACCTGCACACGSCCATCTGCCAGAGAGCTG 189
 Db 33606 CAGGAAACACTCCAGGTAGCTGTGTCTCTCCAGATGACCATTCACCATGAAAAACCG 33549

RESULT 15

US-09-949-016-16518/C
 ; Sequence 16518, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16518
 ; LENGTH: 99749
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-16518

Query Match 14.4%; Score 31.6; DB 4; Length 99749;

Best Local Similarity 54.2%; Pred.No.10; Mismatches 54; Indels 0; Gaps 0;

Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 72 TGCCGAGCCAGCCCTTCATTAAGGCCCTGGGTATGCGCAGCCAGCATGTGCACCTGCCGC 131
 Db 33666 TTCTGAGCAAAACACACACAGCCCAATTGTATACACACCAGAAAGTGGAATGATGGG 33607

QY 132 CGAGACACAAACCCAGCCGAGCATTTGAACCTGCACACGSCCATCTGCCAGAGAGCTG 189
 Db 33606 CAGGAAACACTCCAGGTAGCTGTGTCTCTCCAGATGACCATTCACCATGAAAAACCG 33549

Search completed: September 1, 2005, 07:07:07
 Job time : 73.3315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 1681.14 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-3

Perfect score: 270

Sequence: 1 catggcgacgagggcgacac.....ttccgctactagctagccgc 270

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	212.4	78.7	1181	10	RAT1LGFZ	M84484 Rattus norv
2	212.4	78.7	185148	2	AC136382	AC136382 Rattus no
3	188.4	69.8	1363	10	MM1LGF	X67493 M.musculus
4	188.4	69.8	192843	10	AL607124	AL607124 Mouse DNA
5	98.2	36.4	1569	10	BC013345	BC013345 Mus muscu
6	76.6	28.4	1507	10	AY560836	AY560836 Spemroph
7	76.4	28.3	1500	6	AX401932	AX401932 Sequence
8	76.4	28.3	1500	6	AX827271	AX827271 Sequence
9	76.4	28.3	1500	10	RAT1LGFZ	M58654 Rat IGF bin
10	76.4	28.3	5001	6	AX163782	AX163782 Sequence
11	76.4	28.3	5001	10	RAT1LGFZ	I22979 Rat Insulin
12	65.4	24.2	1510	10	BC078889	BC078889 Rattus no
13	62.6	23.2	3886	9	AY095345	AY095345 Papio anu
14	60.8	22.5	448	11	GG7139	GG7139 IGFBP1 XI.1
15	60.8	22.5	6128	6	AX409747	AX409747 Sequence
16	60.8	22.5	6128	9	HUM1GFBP1A	M74587 Human insul
17	60.8	22.5	6128	11	G19994	G19994 BMS5752 Eri
18	60.8	22.5	6480	9	HUM1GFBP1	M59316 Human insul
19	60.8	22.5	9082	9	AY434089	AY434089 Homo sapi

20	60.8	22.5	69887	9	AC091524
21	59.2	21.9	141539	9	AC146152
22	59.2	21.9	189332	2	AC146117
23	59.2	21.9	200935	2	AC148834
24	51	18.9	194	6	AS7715
25	51	18.9	194	6	AR175909
26	51	18.9	13011	6	E14395
27	51	18.9	13011	6	AX827302
28	51	18.9	13011	10	RN1LPG
29	50.4	18.7	895	9	HUM1GFBP1
30	49	18.1	539	11	G67173
31	44.4	16.4	1480	10	RAT1LGFZ
32	43.2	16.0	2717	10	RAT1LGFZ
33	43.2	16.0	231241	2	AC097039
34	37.8	14.0	7061	6	AX251887
35	37.8	14.0	7061	6	AX345899
36	37.8	14.0	7061	6	AX348680
37	36.8	13.6	348925	1	BX640428
38	36.8	13.6	348925	1	BX640427
39	36.8	13.6	349442	1	BX640447
40	36.6	13.6	415	9	HSPYKINPR
41	36.6	13.6	8409	9	HSPYKINPR
42	36.6	13.6	13322	9	AY316591
43	36.6	13.6	13325	9	AL713999
44	35.6	13.2	300217	1	AE016922
45	35	13.0	511	11	BV184757

ALIGNMENTS

RESULT 1
LOCUS RAT1LGFZ
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION M84484
VERSION M84484.1 GI:204927
KEYWORDS insulin-like growth factor.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 1181)
Unterman,T.G., Lacsos,R.G., McGary,E., Whalen,C. and Goswami,R.G.
Biochem. Biophys. Res. Commun. (1991) In press
COMMENT
Original source text: Rattus norvegicus (strain Sprague-Dawley)
male adult liver DNA.

FEATURES

Location/Qualifiers
1..1181
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

Query Match
Best local similarity 99.5%; Pred. No. 36-52; DB 10; Length 1181;
Matches 213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 52 TCACAGCAAAACAACTTATTGAAACAGGAGATCTTACAGCGTCCCTGCAATCA 111
DB 823 TCACAGCAAAACAACTTATTGAAACAGGAGATCTTACAGCGTCCCTGCAATCA 882
QY 112 TTAACCGTGTGTCGCGAGCCAGCCCTTCATTAAGCCCTGGTATAGCCAGCAGCATGCT 171
DB 883 TTAACCGTGTGTCGCGAGCCAGCCCTTCATTAAGCCCTGGTATAGCCAGCAGCATGCT 942
QY 172 CCATGCGCGCGGAGACAAACAGCAGCATTTGAACATGACACGCGCATCTGCC 231

Db	943	CCACTGCCCCGAGAGACAAACCCAGGAGCATTTGACACAGCGCATCTGCC	1002
Qy	232	AGAGAGCTGTGACCACTTCGGCTACTAGCTA	265
Db	1003	AGAGAGCTGTGACCACTTCGGCTACTATCTA	1036
RESULT 2			
LOCUS	AC136382	185148 bp	DNA linear HMG 01-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS		
ACCESSION	AC136382		
VERSION	AC136382.1	GI:24462257	
KEYWORDS	HTG: HMG_PHASE1		
SOURCE	Rattus norvegicus	(Norway rat)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 185148)		
AUTHORS	Muzny,D,Marie., Metzker,M,lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsdorfs,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,P., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Becotto,M., Eugene,C., Evans,C.A., Faller,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Haml,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Huylk,S., Hume,J., Idlesbit,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., Khan,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshaw,L., Louisedge,H., Lozano,R.U., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojce,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saverly,G., Scherier,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,A., Tingey,A., Trejos,Z., Uemari,K., Valse,R., Vera,V., Villasana,D., Waltron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczky,R., Wooten,H., Wroley,K., Wright,D., Wright,R., Wu,J., Yakub,X., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 185148)		
AUTHORS	Rat Genome Sequencing Consortium.		

TITLE	Direct Submission
JOURNAL	Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	<p>Genome Center</p> <p>Center: Baylor College of Medicine</p> <p>Center code: BCM</p> <p>Web site: http://www.hgsc.bcm.tmc.edu/</p> <p>Contact: hgsc-help@bcm.tmc.edu</p> <p>Center project name: KDGS</p> <p>Center clone name: CH230-97018</p> <p>Sequencing vector: Plasmid</p> <p>Chemistry: Dye-terminator Big Dye 100% of reads</p> <p>Assembly program: Phrap, version 0.990329</p> <p>Consensus quality: 135613 bases at least Q40</p> <p>Consensus quality: 140849 bases at least Q30</p> <p>Consensus quality: 145680 bases at least Q20</p> <p>*****</p> <p>* NOTE: Estimated insert size may differ from sequence length</p> <p>* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).</p> <p>* NOTE: This is a 'working draft' sequence. It currently</p> <p>* consists of 63 contigs. The true order of the pieces</p> <p>* is not known and their order in this sequence record is</p> <p>* arbitrary. Gaps between the contigs are represented as</p> <p>* runs of N, but the exact sizes of the gaps are unknown.</p> <p>* This record will be updated with the finished sequence</p> <p>* as soon as it is available and the accession number will</p> <p>* be preserved.</p> <p>1</p> <p>1011 1110: contig of 1010 bp in length</p> <p>1111 1110: gap of unknown length</p> <p>2609 2708: gap of unknown length</p> <p>2709 3915: contig of 1207 bp in length</p> <p>3916 4015: gap of unknown length</p> <p>4016 5547: contig of 1532 bp in length</p> <p>5548 5647: gap of unknown length</p> <p>5648 7120: contig of 1472 bp in length</p> <p>7120 7220: gap of unknown length</p> <p>8885: contig of 1666 bp in length</p> <p>8886 10258: gap of unknown length</p> <p>10259 10358: contig of 1273 bp in length</p> <p>10359 11626: gap of unknown length</p> <p>11627 11726: contig of 1268 bp in length</p> <p>11727 13488: gap of unknown length</p> <p>13489 15122: contig of 1762 bp in length</p> <p>15123 15222: gap of unknown length</p> <p>15223 16946: contig of 1534 bp in length</p> <p>16947 17046: gap of unknown length</p> <p>17047 18164: contig of 1118 bp in length</p> <p>18165 18264: gap of unknown length</p> <p>18265 19678: contig of 1414 bp in length</p> <p>19679 19778: gap of unknown length</p> <p>19779 21687: contig of 1909 bp in length</p> <p>21688 21787: gap of unknown length</p> <p>21788 22892: contig of 1105 bp in length</p> <p>22893 24336: gap of unknown length</p> <p>24337 24437: contig of 1344 bp in length</p> <p>24438 24436: gap of unknown length</p> <p>24437 26617: gap of unknown length</p> <p>26618 26717: contig of 2181 bp in length</p> <p>26718 28577: gap of unknown length</p> <p>28578 28677: contig of 1860 bp in length</p> <p>28679 30823: gap of unknown length</p> <p>30824 30923: contig of 2146 bp in length</p> <p>30924 3266: gap of unknown length</p> <p>3267 32366: contig of 1343 bp in length</p> <p>32367 33894: gap of unknown length</p> <p>33895 33994: contig of 1528 bp in length</p> <p>33995 35373: gap of unknown length</p> <p>35374: contig of 1379 bp in length</p>

35374 35473: gap of unknown length
35474 37295: contig of 1822 bp in length
37296 37395: gap of unknown length
37396 38420: contig of 1025 bp in length
38421 38520: gap of unknown length
38521 39974: contig of 1454 bp in length
39975 40074: gap of unknown length
40075 42283: contig of 2209 bp in length
42284 42383: gap of unknown length
42384 43721: contig of 1338 bp in length
43722 46538: contig of 2717 bp in length
46539 46638: gap of unknown length
46639 48621: contig of 1983 bp in length
48622 48721: gap of unknown length
48722 50602: contig of 1881 bp in length
50603 50702: gap of unknown length
50703 53841: contig of 3139 bp in length
53842 53941: gap of unknown length
53942 56195: contig of 2254 bp in length
56196 56295: gap of unknown length
56296 58949: contig of 2654 bp in length
58950 61032: contig of 1983 bp in length
61033 61132: gap of unknown length
61133 62942: contig of 1810 bp in length
62943 63042: gap of unknown length
63043 66045: contig of 3003 bp in length
66046 66145: gap of unknown length
66146 68197: contig of 2052 bp in length
68198 68297: gap of unknown length
68298 72235: contig of 3938 bp in length
72236 72335: gap of unknown length
72336 75803: contig of 3468 bp in length
75804 75903: gap of unknown length
75904 78355: contig of 2452 bp in length
78356 78455: gap of unknown length
78456 81458: contig of 3003 bp in length
81459 81558: gap of unknown length
81559 84673: contig of 3115 bp in length
84674 84773: gap of unknown length
84774 87739: contig of 2966 bp in length
87740 87839: gap of unknown length
87840 91535: contig of 3636 bp in length
91536 91635: gap of unknown length
91636 95522: contig of 3887 bp in length
95523 95622: gap of unknown length
95623 98730: contig of 3108 bp in length
98731 98830: gap of unknown length
98831 101618: contig of 2768 bp in length
101619 101718: gap of unknown length
101719 105535: contig of 3817 bp in length
105536 105635: gap of unknown length
105636 109393: contig of 3758 bp in length
109394 109493: gap of unknown length
109494 113564: contig of 4071 bp in length
113565 113664: gap of unknown length
113666 117630: contig of 3966 bp in length
117631 117730: gap of unknown length
117731 122502: contig of 4772 bp in length
122503 122602: gap of unknown length
122603 126644: contig of 4042 bp in length
126645 126744: gap of unknown length

Query Match 78.7%; Score 212.4; DB 2; Length 185148;
Best Local Similarity 99.5%; Pred. No. 3e-52; Indels 0; Gaps 0;
Matches 213; Conservative 0; Mismatches 1;

QY 52 TCACAGCAAAACAACTTATTTTGAACACGGGAGATCTAGACGCTGCTGACAAATCA 111
DB 99253 TCACAGCAAAACAACTTATTTTGAACACGGGAGATCTAGACGCTGCTGACAAATCA 99312
QY 112 TTAACCGCTGCTGCGAGCCAGCCCTTCAATAAGGCGCTCGATATGGCCAGCAGATGCT 171

DB 99313 TTAACCGCTGCTGCGAGCCAGCCCTTCAATAAGGCGCTCGATATGGCCAGCAGATGCT 99372
QY 172 CCACCTGCGCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATCTGCC 231
DB 99373 CCACCTGCGCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATCTGCC 99432
QY 232 AGAGAGCTGTGACCACTTCCGCTACTAGCTA 265
DB 99433 AGAGAGCTGTGACCACTTCCGCTACTAGCTA 99466

RESULT 3
MWITGF
LOCUS MWITGF 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V., Mohn,K.L., Bucan,M. and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression during liver regeneration
JOURNAL Hepatology 19 (3), 656-665 (1994)
MEDLINE 94164648
PUBMED 7509771
REFERENCE
AUTHORS Mohn,K.L., Waddell,J.R. and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential insulin-responsive sequence and conservation of all intron/exon boundaries
JOURNAL Nucleic Acids Res.
REFERENCE
AUTHORS Taub,R.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard Hughes Medical Institute, Clinical Research Bldg, Room 475, 422 Curie Boulevard, Philadelphia, PA 19104-6145, USA

FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus sequences"
764..768
792..1336
/number=1
792
/note="mRNA cap site"
964..>1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org

contact: amadan@systemsbio.org

Anup Madan, Jessika Fahy, Erin Helton, Mark Keltman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN.ac: http://image.lnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1.1569

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:14075 IMAGE:4161889"

/tissue_type="liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1.1569

/gene="Igfbp1"

/note="Synonym: IGFBP-1"

/db_xref="LOCUSID:16006"

/db_xref="MGI:96436"

204.1022

/gene="Igfbp1"

/codon_start=1

/product="insulin-like growth factor binding protein 1"

/protein_id="AAH13345.1"

/db_xref="GI:15426483"

/db_xref="MGI:16006"

/db_xref="MGI:96436"

/translation="MPEPLTVISWPELILSPQIGVAAAGAPQPMHCAPICTARLGLCP

PVPASCPESRPAGCGCCCTCALPMGAGVATATACAGLSCRALPGRRPLHALTRG

QGACVPEPAATSTLSSQHEAKAASVADLSPEMTESQLDSEHLMASPED

QPLMNAISTYSKRAEMADIKKWKQCRRELYVLERLAAQKAGBEIKFYLPN

CNKGFSYHSKQETSIDGBALCMVCVPMGSKIPGSLIETRGDPNCHQYFNVN"

ORIGIN

Query Match 36.4%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 3.2e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 155 TGGCCAGCAGCATGCTCCAGCCGCGAGACAAACCGAGCATTTGAACATG 214
|||
DB 20 TGGGCGAGCATGCTCCAGCCGCGAGACAAACCGAGCATTTGAACATG 79
|||
QY 215 CACAGCGGCATCTGCCAGAGAGCTGTGACCACTTCGGTACTAGCTA 265
|||
DB 80 CACAGCGGCCTCTGCCAGAGAGCTGTGACCACTTCGGTACTAGCTA 130
|||

RESULT 6 1507 bp mRNA linear ROD 22-MAR-2004
AY560836
LOCUS
DEFINITION
Spermophilus tridecemlineatus insulin-like growth factor binding
protein 1 (IGFBP1) mRNA, complete cds.
ACCESSION
AY560836
VERSION
AY560836.1 GI:45505308
KEYWORDS
Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
Spermophilus tridecemlineatus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Murinae;
Spermophilus.
1 (bases 1 to 1507)
REFERENCE
Li,Y., Klimanis,D. and Hallenbeck,J.M.
Cloning and characterization of insulin-like growth factor binding
protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1507)
AUTHORS Li,Y., Klimanis,D. and Hallenbeck,J.M.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
Bethesda, MD 20892, USA

FEATURES

source

1.1507

/organism="Spermophilus tridecemlineatus"

/mol_type="mRNA"

/db_xref="taxon:43179"

1.1507

/gene="IGFBP1"

689.1507

/gene="IGFBP1"

/codon_start=1

/product="insulin-like growth factor binding protein 1"

/protein_id="AAS67029.1"

/db_xref="GI:45505309"

/translation="MPEVPAAGLMPFLILLAVQSTVASSTQPMHCAPICTARLGLCP

PVPSCPESLSPAGCGCCPMCALPLGACGVATATATAGLSCRALPGRRPLHALTRG

QGACVPEPATPTASGLSIEKEAKASVPRVPESEKEMTEBQLDSFHLMASSSD

QPLMNAISTYSKRAEMADIKKWKQCRRELYVLERLAAQKAGBEIKFYLPN

CNKGFSYHSKQETSIDGBALCMVCVPMGSKRIPGSLIETRGDPNCHQYFNVN"

ORIGIN

Query Match 28.4%; Score 76.6; DB 10; Length 1507;
Best Local Similarity 67.6%; Pred. No. 8.8e-12;
Matches 144; Conservative 0; Mismatches 54; Indels 15; Gaps 2;

QY 51 ATCAAGCAAAACAACTATTATTGAACAGGGATCTAGCAGCGCCCTGCAATC 110
|||
DB 429 AGCAGAGCAAAACAACTATTATTGAACAGGGATCTAGCAGCGCCCTGCAATC 488
|||
QY 111 ATTA-----CCGCTGCGCGAGCCAGCCCTCATPAGSCCTGGGTATGGCCA 160
|||
DB 489 ATTAATCTTCTAGTCAAGTGAAGAACCGGCCCTTATTAAGCAAGGCTGACCCA 548
|||
QY 161 GCCAGCATGTCCTCACTGCGCGCGAGACAAACCGAGCATTTGAACATGACACG 220
|||
DB 549 GCAAGCATGTCGCACTGCCAGCGGAGCTCAGAAATTGACACCTATCA-----GCCACT 603
|||
QY 221 GCCATGCGCCAGAGAGCTGTGACCAACCACTTC 253
|||
DB 604 GCCAGCTGCCAGAGAGCATTTGACCACTGTCC 636
|||

RESULT 7
AX401932 1500 bp DNA linear PAT 06-JUN-2002
LOCUS
DEFINITION
Sequence 1608 from Patent WO0210453.
ACCESSION
AX401932
VERSION
AX401932.1 GI:21338112
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE
Mendrick,D., Porter,M.W., Johnson,K.R., Caetle,A.L. and
Elschoff,M.R.
Molecular toxicology modeling
Patent: WO 0210453-A 1608 07-FEB-2002;
JOURNAL
Gene Logic, Inc. (US)
FEATURES
source
1.1500
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

```

Query Match          28.3%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
      |||||||
      8 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67

DB      248 CACTTCGGCTACTAGCTA 265
      |||||||
      68 CACTTCGGCTACTAGCTA 85

RESULT 8
AX827271      1500 bp      DNA      linear      PAT 12-DEC-2003
LOCUS      AX827271
DEFINITION      Sequence 5 from Patent EP1344834.
ACCESSION      AX827271
VERSION      AX827271.1 GI:39837360
KEYWORDS
SOURCE
ORGANISM      Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS      Boess, F., Suter-Dick, L. and Wolf, D.
TITLE      Methods for the toxicity prediction of a compound
JOURNAL      Patent: EP 1344834-A 5 17-SEP-2003;
F. HOFMANN-LA ROCHE AG (CH)
FEATURES
source      1. 1500
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          28.3%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
      |||||||
      8 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67

DB      248 CACTTCGGCTACTAGCTA 265
      |||||||
      68 CACTTCGGCTACTAGCTA 85

RESULT 9
RATGFB      1500 bp      mRNA      linear      ROD 27-APR-1993
DEFINITION      Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.
ACCESSION      M58634
VERSION      M58634.1 GI:204732
KEYWORDS      IGF binding protein-1.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS      Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE      The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL      Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE      91141487
PUBMED      1705004
COMMENT      Original source text: Rat, cDNA to mRNA.
FEATURES
source      Location/Qualifiers
              1. 1500
              /organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="Fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AAA41380.1"
/db_xref="GI:204733"
/translation="MPEFLTVSWPEFLILSFQYRVVAGADQPMHCAPCTAERLELCP
PVPAACEIISRPAGCGCCPTCALPIGAACGVATACAGLSICALPGEPRLHLTRG
QGACVLEPAAPATSLSGSOSHEBAKAAVASDELAESPTEEQLDLSFHLMAFSRED
OPIIMNASTYSSMRAREITDLEKMKPECORELYKVERLAAOQKGDLEYKPYLPN
CNKGFYHSKOCERSLDGEALCMVCYVPMSKRIPLSLERTDGDNCHQYFVQN"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match          28.3%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
      |||||||
      8 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67

DB      248 CACTTCGGCTACTAGCTA 265
      |||||||
      68 CACTTCGGCTACTAGCTA 85

RESULT 10
AX163782      5001 bp      DNA      linear      PAT 22-JUN-2001
LOCUS      AX163782
DEFINITION      Sequence 46 from Patent WO0138579.
ACCESSION      AX163782
VERSION      AX163782.1 GI:14544878
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS      Gould-Rothberg, B.E., Dipippo, V.A., Rameah, T.M. and Gerwein, R.W.
TITLE      Method of identifying toxic agents using msal-d-induced differential
JOURNAL      gene expression in liver
PUBMED      Patent: WO 0138579-A 46 31-MAY-2001;
Curagen Corporation (US)
FEATURES
source      Location/Qualifiers
              1. 5001
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          28.3%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 1e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
      |||||||
      1 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 60

DB      248 CACTTCGGCTACTAGCTA 265
      |||||||
      61 CACTTCGGCTACTAGCTA 78

RESULT 11
RATGFB

```

LOCUS RATTGFA 5001 bp DNA linear ROD 30-NOV-1995
 DEFINITION Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,
 complete cds.
 ACCESSION L22979
 VERSION L22979.1 GI:1098472
 KEYWORDS insulin-like growth factor binding protein-1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 5001)
 REFERENCE 1
 AUTHORS Iacson, R., Oehler, D., Yang, B., Goswami, R. and Unterman, T.
 TITLE Bideoxy sequencing and structural analysis of the rat insulin-like
 growth factor binding protein-1 gene
 JOURNAL Biochim. Biophys. Acta 1218 (1), 95-98 (1994)
 MEDLINE 94250701
 PUBMED 7514892
 COMMENT On Nov 30, 1995 this sequence version replaced gi:185167.
 Original source text: Rattus norvegicus (strain Sprague-Dawley)
 DNA.

FEATURES
 source
 1. 5001
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /sex="male"
 /issue_type="11ver"
 /dev_stage="adult"
 /cissue_11b="Clontech EMBL-3 Sp6/T7"
 60..525
 /number=1
 exon
 5'UTR
 60..152
 153..4221
 /gene="IGFBP-1"
 join(153..525,1850..2034,2874..3002,4090..4221)
 CDS
 /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein"
 /protein_id="AA82581.1"
 /db_xref="GI:1098473"
 /translation="MPEFLTVVSWPFLILSFQVAVVAGAPQPMHCAPCTARLELCP
 VPASCPRTSRAGCGCCCTCALPLGAAAGVATAACAGLSRAALPGEPRLHALTRG
 OGACVDEPAAPATSSLSGSQHEAKAAVSEDLAESPEMTESQLDGFHLMAPSRD
 OPTLWNAISTYSKSMARREITDLKKWKEPQRLVRLAAQOKADDEIKFPLPN
 CNKNGFYHSKCKETSLDGEAGLCMCVPMGSKIKPGSLRETRDPRCHQYFNVQN"
 526..1849
 /gene="IGFBP-1"
 /number=1
 exon
 1850..2034
 /gene="IGFBP-1"
 /number=2
 intron
 2035..2873
 /gene="IGFBP-1"
 /number=2
 exon
 2874..3002
 /gene="IGFBP-1"
 /number=3
 intron
 3003..4089
 /gene="IGFBP-1"
 /number=3
 exon
 4090..4743
 /number=4
 4744..4743
 ORIGIN
 3'UTR
 4222..4743

Query Match 28.3%; Score 76.4; DB 10; Length 5001;
 Best Local Similarity 98.7%; Pred. No. 1e-11;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 188 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 247
 |||||||

Db 1 CACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 60
 248 CACTTCGGCTACTAGCTA 265
 |||||||
 Db 61 CACTTCGGCTACTAGCTA 78
 |||||||
 RESULT 12
 BC078889
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 1 (bases 1 to 1510)
 REFERENCE 1
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshllyuk, S.,
 Carninci, P., Prange, C., Raha, S.S., Loguellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Holly, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1510)
 DIRECTOR MGC Project.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 COMMENT

REMARK
 COMMENT
 Contact: MGC help desk
 Email: cgaabp@remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www.shgc.stanford.edu
 Contact: (Dickson, Mark) mcdpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRAK Plate: 184 Row: e Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6981079.
 Location/Qualifiers
 1. 1510
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"

/clone="MGC:93595 IMAGE:7129185"
 /tissue_type="kidney", rat (Brown Norway)"
 /clone_lib="NIH_MGC_236"
 /lab_host="DH108"
 /note="vector: pExpress1"
 1. .1510
 /gene="Igf1bp1"
 /note="synonyms: IGFBP, IGFBP25, IBP1"
 /db_xref="locusID:25685"
 /db_xref="RATMAP:44422"
 /db_xref="RGD:2872"
 142. .960
 /gene="Igf1bp1"
 /codon_start=1
 /product="insulin-like growth factor binding protein 1"
 /protein_id="AAH78889.1"
 /db_xref="GI:50927647"
 /db_xref="locusID:25685"
 /db_xref="RATMAP:44422"
 /db_xref="RGD:2872"
 /translation="MPEFLTVVSPFLILSFQVAVGAPDPWHCAPCTARBLCTC
 PVPCAPSPISPPAGCCPTCALPGACGVATACACGSLRGLRGPRLHATRR
 OGACVLRPAAPRTSSLSQHEEAAVAASEBLAESPMTEBQLDSFHLMAPRE
 OPTLNAISTYSMPAREITDLKWKKEPCQRLYVLRLLAAQQADELYKFLP
 CNKNGFVHSHKQETSLLDEAGLWCVPWMSGKRIPGSLDREDPHCQYFNVVN"

Query Match 24.2%; Score 65.4; DB 10; Length 1510;
 Best Local Similarity 98.5%; Pred. No. 1.9e-08;
 Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 199 CGAGCATTTGAACACTGCACACGGCCATCTGCCACAGAGCTGTGACCACTTCGGCTA 258
 1 CGAGCATTTGAACACTGCACACGGCCATCTGCCACAGAGCTGTGACCACTTCGGCTA 60
 |||||
 Oy 259 CTAGCTA 265
 |||||
 Db 61 CTATCTA 67

RESULT 13
 LOCUS AY095345 3886 bp DNA linear PRI 30-DEC-2002
 DEFINITION Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
 ACCESSION AY095345
 VERSION AY095345.1 GI:20853764
 KEYWORDS
 SOURCE
 ORGANISM Papio anubis (olive baboon)
 Papio anubis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Papio.
 1 (bases 1 to 3886)
 Kim,J.J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
 Jaffe,R.C., Fazleabas,A.T. and Utermann,T.G.
 Regulation of insulin-like growth factor binding protein-1 promoter
 activity by FHR and HMOX10 in primate endometrial cells
 Biol. Reprod. 68 (1), 24-30 (2003)
 2 (bases 1 to 3886)
 Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
 Direct Submission
 Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
 University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
 USA

FEATURES
 source
 1. .3886
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /db_xref="taxon:9555"
 <3655. .>3886
 /gene="IGFBP-1"

gene

DB	QY	Query Match	Best Local Similarity	Matches	135;	Conservative	0;	Indels	17;	Gaps	2;
3539	CACCTCCCAAGAGCACTGGCCACCGCTCC	3569									
QY	223	CATCTGCCAAGAGCTGTGACCACTTC	253								
DB	3486	AAGCATGGGGCCAGCGCATCCCATCCAGCAAGC-----ATCTGCGCGCGCGCGCGCG	3538								
QY	163	CAGCATGGTCCACTGCGCGCGAGACACAACCAAGAGAGCATTGAACATGCACACGCG	222								
DB	3426	TAACTCTCGTGTGCAAGTGGCGCGCGCTGTGTCCCTTTATTAAGGGCGCGCGCTGTGTCCAGC	3485								
QY	113	TAAACCC-----GTGCTGCCGAGCGACGCCCTTCATTAAGGCGCTGGGGTATGCGCAGC	162								
DB	3366	CACATGCAAAACAACCTTATTTTGAACACTGAGTCTCTAGAGTGGCCGGCGTGCACATCAT	3425								
QY	53	CACAAGCAAAACAACCTTATTTTGAACACGGGGATCTCTAGACACCGCTCCCTGCACATCAT	112								

RESULT 14					
LOCUS	G67139	448 bp	DNA	linear	STS 18-SEP-2000
DEFINITION	IGBPI X1.1 Human Homo sapiens STS genomic, sequence tagged site.				
ACCESSION	G67139				
VERSION	G67139.1	GI:10186730			
KEYWORDS	STS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 448)				
AUTHORS	Cox,D.G., Boillot,C. and Canzian,F.				
TITLE	Genome Survey				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Federico Canzian Genome Analysis Group International Agency for Research on Cancer 150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France Tel.: +33-4-72738698 Fax: +33-4-72738388 Email: canzian@iarc.fr Primer A: TGCATCAGCAACAAACTT Primer B: GGGTGACCTCCGAGCAGC STS size: 448 Protocol: Template: 50 ng Primer: 4 uM each dNTPs: 2 mM each MgCl2: 1.5-2.5 mM Tag: 0.05 units Total Vol: 25 uL Buffer: MgCl2: 1.5-2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.4.				

FEATURES
source 1. 448
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind 1. 448
primer_bind 1. 20
complement(431. 448)
ORIGIN

Query Match 22.5%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 4.5e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCTGCAATCAT 112
DB 3 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGGCGCTGCCAATCAT 62
QY 113 TAAACC-----GTGCTGCCGAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 162
DB 63 TAACTCTCTGTGCAAGTGCGCGGCTGTGCTTTATTAAGTGCGCGCTGTGTCAGC 122
QY 163 CAGCATGTCTCACTGCGCGCGAGACAAACCGAGCATTTGAACACTGC-ACACGG 221
DB 123 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTGCCGCGCGCGCG 174
QY 222 CCATCTGCCGAGAGAGCTGTGACCACTTC 253
DB 175 CCACCTCCGAGAGAGACTGGCCACCGCTCC 206

RESULT 15
LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2194 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:21442452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
Alvarez, C., Horne, D., Perez-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source 1. 6128
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN
Query Match 22.5%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 4.5e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCTGCAATCAT 112
DB 473 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGGCGCTGCCAATCAT 532
QY 113 TAAACC-----GTGCTGCCGAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 162
DB 533 TAACTCTCTGTGCAAGTGCGCGGCTGTGCTTTATTAAGTGCGCGCTGTGTCAGC 592
QY 163 CAGCATGTCTCACTGCGCGCGAGACAAACCGAGCATTTGAACACTGC-ACACGG 221
DB 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTGCCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCACTTC 253

DB 645 CCACCTCCGAGAGAGACTGGCCACCGCTCC 676

Search completed: September 1, 2005, 03:26:49
Job time : 1685.14 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 / Search time 2052.72 Seconds
(without alignments)
5006.706 Million cell updates/sec

Title: US-09-972-916B-3

Perfect score: 270
Sequence: 1 catggggcgcacggggcgcaccc.....ttccgctactagctagccgc 270

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hcc: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.2	36.4	1013	2	BF236974 602026605
2	87.8	32.5	310	6	BY783538 BY783538
3	87.8	32.5	331	6	BY794229 BY794229
4	87.8	32.5	354	6	BY771317 BY771317
5	86.8	32.1	401	1	AT785818 AT785818
6	86.8	32.1	480	1	AT196314 AT196314
7	86.8	32.1	706	1	AT153016 AT153016
8	86.8	32.1	765	1	AT098594 AT098594
9	86.8	32.1	785	1	AT153013 AT153013
10	86.8	32.1	799	1	AT529939 AT529939
11	86.8	32.1	811	6	CA478518 AGENCOURT
12	86.8	32.1	848	1	AT790802 AT790802
13	86.8	32.1	852	1	AT528304 AT528304
14	85.2	31.6	605	1	AT196154 AT196154
15	82.2	30.4	380	1	AT785039 AT785039
16	81.4	30.1	846	7	COS73026 AGENCOURT
17	80.4	29.8	692	7	CV127049 AGENCOURT
18	79.4	29.4	713	7	COS60662 AGENCOURT
19	75.4	27.9	748	7	COS75629 AGENCOURT
20	73.4	27.3	694	7	CV117001 AGENCOURT
21	72.4	27.1	332	7	W30013 W30013
22	72.4	26.8	615	7	BB660958 BB660958
23	72.4	26.8	618	6	CD561711 CD561711
24	72.4	26.8	632	1	AT892189 AT892189

25	67.4	25.0	488	1	AA674302 AA674302
26	65.4	24.2	801	7	CK472246 AGENCOURT
27	64.4	23.9	840	7	CK473709 AGENCOURT
28	62.8	23.3	759	1	AA105355 mp37d09.1
29	61.8	22.9	269	2	BB604790 BB604790
30	61.6	22.8	587	2	AW916227 EST347531
31	60.8	22.5	1589	3	CR621807 full-1eng
32	60.8	22.5	1601	3	CR595377 full-1eng
33	59	21.9	545	1	AA060360 m67a12.1
34	43	15.9	659	8	AZ840793 2M0138D02
35	38.8	14.4	698	5	BU684425 UI-CR-ENO
36	37.2	13.8	552	2	AW655007 105511 MA
37	36.8	13.6	428	2	BR829782 MR3-HN006
38	36.8	13.6	849	9	CNS0206E Tetraodon
39	36.6	13.6	692	7	T87256 yd91c06.r1
40	35	13.0	342	9	BX282084 BX282084
41	35	13.0	490	5	AA443181 zxn9h10.1
42	35	13.0	509	1	AA443181 zxn9h10.1
43	35	13.0	532	4	BG383072 60240685
44	35	13.0	533	4	BG388790 602414432
45	35	13.0	561	4	BM772026 K-EST0056

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4161889 5', mRNA sequence.
ACCESSION BF236974
VERSION BF236974.1 GI:11150891
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1013)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence atp: 581.
Location/Qualifiers
1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="TIMGR:4161889"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_L19"
/note="Organ: liver; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."
ORIGIN
Query Match 36.4%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 1.5e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 155 TGGCAGGACGACGATGTCCTGCGCCGCGAGACAAACCGAGGAGATTGAACACTG 214

```

Db      9  TGGGAGCGCAGCATGGTCCACTGCGCGGAGACACACGCCAGGAGATTGAACACTG 68
OY      215 CACAGCGGCATCTGCCCCAGAGAGCTGTATCCACCACTTCCGCTACTACTGTA 265
Db      69  CACAGCGCGCTGTCCCGAGAGAGCTGTATCCACCACTTGCCTACTACTACTTA 119

RESULT 2
BY783538
LOCUS   BY783538
DEFINITION Mus musculus full-length enriched, 17.5 days embryo whole body
ACCESSION BY783538
VERSION   BY783538
KEYWORDS  BY783538.1 GI:39710177
SOURCE    EST
ORGANISM  Mus musculus (house mouse)

REFERENCE
AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
          Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
          Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
          Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
          Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
          Pavan,W., Aldinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
          Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
          Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
          Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
          Targeting a complex transcriptome: the construction of the mouse
          full-length cDNA encyclopedia
          Genome Res. 13 (6B), 1273-1289 (2003)
          22703353
          12819125

TITLE
JOURNAL  Contact: Yoshihide Hayashizaki
MEDLINE  Laboratory for Genome Exploration Research Group, RIKEN Genomic
PUBMED   Sciences Center(GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in Riken
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site (http://genome.gsc.riken.jp/) for
          further details.

FEATURES
source   Location/Qualifiers
          1..310
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="C57BL/6J"
          /db_xref="taxon:10090"
          /clone="U930176D05"
          /tissue_type="whole body"
          /dev_stage="17.5 days embryo"
          /clone_idb="RIKEN full-length enriched, 17.5 days embryo
          whole body"

ORIGIN
Query Match      32.5%; Score 87.8; DB 6; Length 310;
Best Local Similarity 92.9%; Pred. No. 1.6e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      167 ATGTGTCACCTGCGCGGAGACACAACCCAGGAGCATTTGAACATGCACGCGCCATC 226
Db      2  ATGTGTCACCTGCGCGGAGACACAACCCAGGAGCATTTGAACATGCACGCGCCGTC 61
OY      227 TGCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTTA 265

```

```

Db      62  TGCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTTA 100

RESULT 3
BY794229
LOCUS   BY794229
DEFINITION Mus musculus full-length enriched, 17.5 days embryo whole body
ACCESSION BY794229
VERSION   BY794229
KEYWORDS  BY794229.1 GI:39720868
SOURCE    EST
ORGANISM  Mus musculus (house mouse)

REFERENCE
AUTHORS   Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M.,
          Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
          Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watahiki,A.,
          Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
          Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
          Pavan,W., Aldinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
          Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M.,
          Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
          Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
          Targeting a complex transcriptome: the construction of the mouse
          full-length cDNA encyclopedia
          Genome Res. 13 (6B), 1273-1289 (2003)
          22703353
          12819125

TITLE
JOURNAL  Contact: Yoshihide Hayashizaki
MEDLINE  Laboratory for Genome Exploration Research Group, RIKEN Genomic
PUBMED   Sciences Center(GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/
          cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in Riken
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site (http://genome.gsc.riken.jp/) for
          further details.

FEATURES
source   Location/Qualifiers
          1..331
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="C57BL/6J"
          /db_xref="taxon:10090"
          /clone="U930292H24"
          /tissue_type="whole body"
          /dev_stage="17.5 days embryo"
          /clone_idb="RIKEN full-length enriched, 17.5 days embryo
          whole body"

ORIGIN
Query Match      32.5%; Score 87.8; DB 6; Length 331;
Best Local Similarity 92.9%; Pred. No. 1.6e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      167 ATGTGTCACCTGCGCGGAGACACAACCCAGGAGCATTTGAACATGCACGCGCCATC 226
Db      2  ATGTGTCACCTGCGCGGAGACACAACCCAGGAGCATTTGAACATGCACGCGCGTC 61
OY      227 TGCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTTA 265
Db      62  TGCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTTA 100

```

```

RESULT 4
BY71317

```


VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
A1785818.1	GI:5333534							
EST	Mus musculus (house mouse)							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.							
	Mus musculus							
	1 (bases 1 to 401)							
	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepec,M., Thaising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Riter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.							
	The WashU-NCI Mouse EST Project 1999							
	Unpublished (1999)							
	Other ESTs: u178h05.x1							
	Contact: Marra M/WashU-NCI Mouse EST Project 1999							
	Washington University School of Medicine							
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA							
	Tel: 314 286 1800							
	Fax: 314 286 1810							
	Email: mouseest@watsn.wustl.edu							
	This clone is available royalty-free through LNL; contact the							
	IMAGE Consortium (info@image.lnl.gov) for further information.							
	MGI:972893							
	Seq primer: custom primer used							
	High quality sequence stop: 126.							
	Location/Qualifiers							
	1. 401							
	/organism="Mus musculus"							
	/mol_type="mRNA"							
	/strain="C57BL"							
	/db_xref="taxon:10090"							
	/clone="IMAGE:1888569"							
	/sex="female"							
	/dev_stage="adult"							
	/lab_host="DH10B"							
	/clone_1lb="Sugano mouse liver m1a"							
	/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer							
	[ATGAGCCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was							
	ligated to a DraIII adaptor (TGTGGCTACTG), digested							
	and cloned into distinct DraIII sites of the pME18S-FL3							
	vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should							
	be used to isolate the cDNA insert. Size selection was							
	performed to exclude fragments <1.5kb. Library							
	constructed by Dr. Sumio Sugano (University of Tokyo							
	Institute of Medical Science). Custom primers for							
	sequencing: 5' end primer CTCTGTCTCTTAAAGCTGCG and 3' end							
	primer CGACTGTGACGCTCGACACA."							
ORIGIN								
Query Match	32.1%;	Score 86.8;	DB 1;	Length 401;				
Best Local Similarity	92.9%;	Pred. No. 3.4e-15;						
Matches 91;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;				
Oy	168	TGTCCTCCTCGCCGCGAGACACAAACCGAGCATTTGACACCTGCACAGCGCCATCT	227					
Db	1	TGGTCACCTGCGCGGAGACACACACCCAGCGACATTTAAACACTGCACACGCGCTCT	60					
Oy	228	GCCGAGAGACTGTGACACCACTTCCTCGCTACTAGCTA	265					
Db	61	GCCGAGAGACTGTGACACCACTTCCTCGCTACTAGCTA	98					
RESULT 6								
LOCUS	A1196314							
DEFINITION	480 bp	mRNA	linear	EST 14-OCT-1998				
	u171a07.v1	Sugano mouse liver m1a	Mus musculus	cDNA clone				
	IMAGE:1887828	5' similar to gb:M59316	rna1	INSULIN-LIKE GROWTH				
	FACTOR BINDING PROTEIN 1	PRECURSOR (HUMAN);	gb:X81579	M. musculus				
	mRNA for insulin-like growth factor binding (MOUSE);							
	sequence.							

ACCESSION A1196314 GI:3748920
VERSION A1196314.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 480)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellendberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Watson, R. and Waterston, R.
TITLE The Mashu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Mashu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:973152
Seq primer: custom primer used
High quality sequence stop: 375.
Location/Qualifiers
1..480
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887828"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer (ATGTGGCCCTTTTCTTTTCTT); double-stranded cDNA was ligated to a DraIII adaptor (TTGTGGCCCTACTG), digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGTGAGCTGAGCACA."

ORIGIN
Query Match 32.1%; Score 86.8; DB 1; Length 480;
Best Local Similarity 92.9%; Pred. No. 3.6e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCCCCCGAGACACAAACCAGGAGATTGAACACTGCACAGCGCATCT 227
|||||
DB 1 TGGTCCACTGCCCCCGAGACACACACCAGGAGATTGAACACTGCACAGCGCATCT 60
|||||

QY 228 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 265
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 98
|||||

RESULT 7
A1530146 706 bp mRNA linear EST 18-MAR-1999
LOCUS A1530146
DEFINITION u189f09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1889609 5' similar to gb:W5316 rat insulin-like growth
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
ACCESSION A1530146 GI:4444281
VERSION A1530146.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 706)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The Mashu-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:973933
Seq primer: custom primer used
High quality sequence stop: 479.
Location/Qualifiers
1..706
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889609"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer (ATGTGGCCCTTTTCTTTTCTT); double-stranded cDNA was ligated to a DraIII adaptor (TTGTGGCCCTACTG), digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGTGAGCTGAGCACA."

ORIGIN
Query Match 32.1%; Score 86.8; DB 1; Length 706;
Best Local Similarity 92.9%; Pred. No. 3.9e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCCCCCGAGACACAAACCAGGAGATTGAACACTGCACAGCGCATCT 227
|||||
DB 1 TGGTCCACTGCCCCCGAGACACACACCAGGAGATTGAACACTGCACAGCGCATCT 60
|||||

QY 228 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 265
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCCTCGCTACTAGCTA 98
|||||

RESULT 8
A1098594 765 bp mRNA linear EST 20-AUG-1998
LOCUS A1098594
DEFINITION u218e07.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1481988 5' similar to gb:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION A1098594

VERSION A1098594.1 GI:3448119
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 765)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930344
 Seq primer: custom primer used
 High quality sequence stop: 395.
 Location/Qualifiers
 source
 1..765
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1481988"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 765;
 Best Local Similarity 92.9%; Pred. No. 4e-15; 7; Indels 0; Gaps 0;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGTTCACCTGCCCCCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCACT 227
 DB 1 TGTTCACCTGCCCCCGAGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCGCT 60
 QY 228 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 9
 LOCUS A1530313 785 bp mRNA linear EST 18-MAR-1999
 DEFINITION u191f01.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
 VERSION A1530313.1 GI:4444448
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 785)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence stop: 459.
 Location/Qualifiers
 source
 1..785
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889785"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 785;
 Best Local Similarity 92.9%; Pred. No. 4e-15; 7; Indels 0; Gaps 0;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGTTCACCTGCCCCCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCACT 227
 DB 1 TGTTCACCTGCCCCCGAGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCGCT 60
 QY 228 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 10
 LOCUS A1529939 799 bp mRNA linear EST 18-MAR-1999
 DEFINITION u167c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889392 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION	AI529939
VERSION	AI529939.1
KEYWORDS	GI:4444074
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 799)
AUTHORS	Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra W/MashU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:973716 Seq primer: custom primer used High quality sequence stop: 506. Location/Qualifiers 1..799
FEATURES	
SOURCE	

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889392"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone.lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMT18s-Fl3; Site_1: DraIII
(CACGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [GTGTGCGCATGTG], digested
and cloned into distinct DraIII sites of the pMT18s-Fl3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTTCTGCTTAAAGCTGCG and 3' end
primer CGACCTGACGCTCGAGCA."

```

[illegible]

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
	Mus musculus (house mouse)					
	Mus musculus					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
	1 (bases 1 to 811)					
	NIH-MGC http://mgc.nci.nih.gov/ .					
	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: gsapbs-r@mail.nih.gov					
	Tissue Procurement: Bradfield Laboratory					
	cDNA Library Preparation: Mark Bittinger					
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
	DNA Sequencing by: Agencourt Bioscience Corporation					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LNL at:					
	http://image.lnl.gov					
	Plate: LMK00138 row: m column: 23					
	High quality sequence stop: 536.					
FEATURES	Location/Qualifiers					
source	1..811					

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6766752"
/lab_host="IDH10B (phage-resistant)"
/clone_lib="NIH MGC 152"
/note="Organ: Liver; Vector: pDONR201; Site_1: atp2;
Site_2: atp1; cDNA made by oligo-dT with atb2 site and
directionally cloned. Priming sequence:
5'-TTTCGTGAGGCGCGCACCCACTGTGACAGAAAGCGGCTTTTTTTTTTTT
TTTT-3'. Full-length enriched library was constructed
using the Genescreen kit by Invitrogen, library
amplification 16 cycles. Library constructed by Mark
Bitterling in the Bradfield laboratory (McGardle Laboratory
for Cancer Research, University of Wisconsin). Note: this
is a NIH MGC library."

```

[illegible]

COMMENT

Other ESTs: uk28b10.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039

FEATURES

Seq primer: custom primer used
High quality sequence stop: 514.
Location/Qualifiers

1..848

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse kidney mklia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 32.1%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 4e-15;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCGCCCGGAGACACAAACCAGGAGCATTTGAACACTGCACAGCGCATCT 227

DB 1 TGGTCCACTGCGCCCGGAGACACACACCAGGAGCATTTGAACACTGCACAGCGCATCT 60

QY 228 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 265

DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 13
A1528304 852 bp mRNA linear EST 18-MAR-1999

LOCUS A1528304
DEFINITION u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone

IMAGE:1899210.5, similar to gb:MS9316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1528304
VERSION A1528304.1 GI:4442439

KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 852)

REFERENCE

AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Peterson,B., Schellier,T., Gibbons,M., Page,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534

FEATURES

Seq primer: custom primer used
High quality sequence stop: 478.
Location/Qualifiers

1..852

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1899210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 32.1%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 4e-15;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCGCCCGGAGACACAAACCAGGAGCATTTGAACACTGCACAGCGCATCT 227

DB 1 TGGTCCACTGCGCCCGGAGACACACACCAGGAGCATTTGAACACTGCACAGCGCATCT 60

QY 228 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 265

DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998

LOCUS A1196154
DEFINITION u195d08.y1 Sugano mouse liver mla Mus musculus cDNA clone

IMAGE:1887665.5, similar to gb:MS9316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1196154
VERSION A1196154.1 GI:3748760

KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 605)

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellier,B., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HIMI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsen.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:9721987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

```
1..605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME185-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TTTGGCTTACTGG), digested
and cloned into distinct DraIII sites of the pME185-FL3
vector (5' site CACTGTGTG, 3' site CACCATTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer CGACTGCGACTGCGACACA."
```

ORIGIN

Query Match

Best Local Similarity 31.6%; Score 85.2; DB 1; Length 605;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCCCCGCGAGACACAAACCAGGAGCATTTGAACACTGCACACGGCCATCT 227
|||||
DB 1 TGGTCCACTGCCCCGCGAGACACACACCAGGAGCATTTGAACACTGCACACGGTCTCT 60
|||||
QY 228 GCCCAGAGAGCTGTGAACCACTTCGCTACTTACTGCTA 265
|||||
DB 61 GCCCAGAGAGCTGTGAACCACTTCGCTACTTACTTCTA 98
|||||

RESULT 15

AT785039

LOCUS

DEFINITION

u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone.
IMAGE:1888018 5' similar to gp:X81579 M.musculus mRNA for
inulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AT785039 380 bp mRNA linear EST 02-JUL-1999
u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone.
IMAGE:1888018 5' similar to gp:X81579 M.musculus mRNA for
inulin-like growth factor binding (MOUSE);, mRNA sequence.

REFERENCE

AUTHORS

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theisler, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schumk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

TITLE

JOURNAL

Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsen.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

```
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME185-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TTTGGCTTACTGG), digested
and cloned into distinct DraIII sites of the pME185-FL3
vector (5' site CACTGTGTG, 3' site CACCATTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer CGACTGCGACTGCGACACA."
```

ORIGIN

Query Match

Best Local Similarity 30.4%; Score 82.2; DB 1; Length 380;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 171 TCCACTGCCGCCGCGAGACACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCC 230
|||||
DB 1 TCCACTGCCGCCGCGAGACACACACCAGGAGCATTTGAACACTGCATTAACGGCCGTCTGCC 60
|||||
QY 231 CAGAGAGCTGTGAACCACTTCGCTACTTACTGCTA 265
|||||
DB 61 CAGAGAGCTGTGAACCACTTCGCTACTTACTTCTA 95
|||||

Search completed: September 1, 2005, 06:56:48
Job time : 2054.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 786.359 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-3

Perfect score: 270

Sequence: 1 catggcgccagcgaggcgactc.....ttccgcctactagctacgcgc 270

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 327445616 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270	100.0	270	US-09-972-916A-3	Sequence 3, Appli
2	270	100.0	423	US-09-972-916A-6	Sequence 6, Appli
3	219.4	81.3	321	US-09-972-916A-4	Sequence 4, Appli
4	219.4	81.3	372	US-09-972-916A-5	Sequence 5, Appli
5	219	81.1	219	US-09-972-916A-2	Sequence 2, Appli
6	76.4	28.3	1500	US-09-917-800A-1608	Sequence 1608, Ap
7	76.4	28.3	1500	US-10-388-934-5	Sequence 5, Appli

8	76.4	28.3	1500	17	US-10-191-803-73	Sequence 73, Appli
9	76.4	28.3	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
10	60.8	22.5	6128	9	US-09-880-107-2393	Sequence 2393, Ap
11	60.8	22.5	6128	22	US-10-756-149-1484	Sequence 1484, Ap
12	60.8	22.5	9173	22	US-10-893-315-126	Sequence 126, App
13	60.8	22.5	9174	22	US-10-893-315-160	Sequence 160, App
14	51	18.9	51	9	US-09-972-916A-1	Sequence 1, Appli
15	51	18.9	321	9	US-09-972-916A-4	Sequence 4, Appli
16	51	18.9	372	9	US-09-972-916A-5	Sequence 5, Appli
17	51	18.9	13011	17	US-10-388-934-36	Sequence 36, Appli
18	51	18.9	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
19	37.8	14.0	7061	15	US-10-311-455-970	Sequence 970, App
20	37.8	14.0	7061	17	US-10-221-613-148	Sequence 148, App
21	36.6	13.6	31703	17	US-10-085-117-172	Sequence 172, App
22	36.4	13.5	316778	13	US-10-027-632-174961	Sequence 174961, Sequence 174961,
23	36.4	13.5	316778	17	US-10-027-632-174961	Sequence 174961, Sequence 174961,
24	35.2	13.0	1687	20	US-10-739-930-5461	Sequence 5461, Ap
25	35	13.0	769	13	US-10-027-632-164336	Sequence 164336, Sequence 164336,
26	35	13.0	3895	14	US-10-027-632-164336	Sequence 164336, Sequence 164336,
27	35	13.0	3895	14	US-10-011-858A-76	Sequence 76, Appli
28	34	12.6	43979	19	US-10-741-601-5746	Sequence 5746, Appli
29	34	12.6	43979	21	US-10-741-600-17905	Sequence 17905, A
30	33.4	12.4	1806	19	US-10-250-615-33	Sequence 33, Appli
31	33.4	12.4	2667	17	US-10-108-260A-2239	Sequence 2239, Ap
32	32.4	12.0	1389	19	US-10-437-963-79162	Sequence 79162, A
33	32.4	12.0	1584	19	US-10-437-963-79185	Sequence 79185, A
34	32	11.9	76180	19	US-10-332-281-492	Sequence 492, App
35	31.6	11.7	1129	20	US-10-343-903-31	Sequence 31, Appli
36	31.6	11.7	1288	15	US-10-017-161-2017	Sequence 2017, Ap
37	31.6	11.7	1288	17	US-10-292-798-1663	Sequence 1663, Ap
38	31.6	11.7	1348	17	US-10-072-988-404	Sequence 404, App
39	31.6	11.7	1513	18	US-10-072-912-29	Sequence 29, Appli
40	31.6	11.7	1555	10	US-09-764-682-581	Sequence 581, App
41	31.6	11.7	11750	19	US-10-367-094-169	Sequence 169, App
42	31.4	11.6	360	20	US-10-425-115-28332	Sequence 28332, A
43	31.4	11.6	1134	17	US-10-282-122A-33251	Sequence 33251, A
44	31.4	11.6	16554	13	US-10-087-192-391	Sequence 391, App
45	31.2	11.6	668	20	US-10-363-345A-40493	Sequence 40493, A

ALIGNMENTS

RESULT 1
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thiele, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 100.0%; Score 270; DB 9; Length 270;
Best Local Similarity 100.0%; Pred. No. 3, 2e-86;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTGTGAGCCCAAGTATCAACAGCA 60
Db 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTGTGAGCCCAAGTATCAACAGCA 60
Qy 61 AAACAATTATTTTGAACACAGCGGAGTCTTAGACACGCTGCCCTGACATCATTAACCGGT 120


```
Db      61 AAAAACAATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAATCAATTAACCGT 120
Qy      121 GGTGGCAGACCGACCCCTTATTAAGCCCTGGGTATGCGCAAGCATGTGTCCACTGCCCC 180
Db      121 GGTGGCAGACCGACCCCTTATTAAGCCCTGGGTATGCGCAAGCATGTGTCCACTGCCCC 180
Qy      181 GCCGAGACACAAACCCAGGAGCATTTGAACACTGCAACGAGCATCTGCGCAGAGAGCTG 240
Db      181 GCCGAGACACAAACCCAGGAGCATTTGAACACTGCAACGAGCATCTGCGCAGAGAGCTG 240
Qy      241 TGACCAACCACTTCCGCTACTAGTAGCCGC 270
Db      241 TGACCAACCACTTCCGCTACTAGTAGCCGC 270

RESULT 2
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match      100.0%; Score 270; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 3.7e-86;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CATGGCGCAGCGGGGACTCCCGTGGTCTTGAGACTTGAGCCGCCAGTGTATCAGACGA 60
Db      154 CATGGCGCAGCGGGGACTCCCGTGGTCTTGAGACTTGAGCCGCCAGTGTATCAGACGA 213
Qy      61 AAACAACCTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAATCAATTAACCGT 120
Db      214 AAACAACCTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAATCAATTAACCGT 273
Qy      121 GGTGGCAGACCGACCCCTTATTAAGCCCTGGGTATGCGCAAGCATGTGTCCACTGCCCC 180
Db      274 GGTGGCAGACCGACCCCTTATTAAGCCCTGGGTATGCGCAAGCATGTGTCCACTGCCCC 333
Qy      181 GCCGAGACACAAACCCAGGAGCATTTGAACACTGCAACGAGCATCTGCGCAGAGAGCTG 240
Db      334 GCCGAGACACAAACCCAGGAGCATTTGAACACTGCAACGAGCATCTGCGCAGAGAGCTG 393
Qy      241 TGACCAACCACTTCCGCTACTAGTAGCCGC 270
Db      394 TGACCAACCACTTCCGCTACTAGTAGCCGC 423

RESULT 3
US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
```

```
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match      81.3%; Score 219.4; DB 9; Length 321;
Best Local Similarity 99.5%; Pred. No. 4.7e-68;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      50 TATCAACAAGCAAAACAACTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAAT 109
Db      101 TGTCAACAAGCAAAACAACTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAAT 160
Qy      110 CATTAAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGTATGCGCAGCAGCATG 169
Db      161 CATTAAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGTATGCGCAGCAGCATG 220
Qy      170 GTCCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGCAACGCGCATCTGC 229
Db      221 GTCCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGCAACGCGCATCTGC 280
Qy      220 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270
Db      281 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 321

RESULT 4
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match      81.3%; Score 219.4; DB 9; Length 372;
Best Local Similarity 99.5%; Pred. No. 5e-68;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      50 TATCAACAAGCAAAACAACTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAAT 109
Db      152 TGTCAACAAGCAAAACAACTATTATTTGAACACGGGAGTCTTAGACAGCTGCCCTGACAAAT 211
Qy      110 CATTAAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGTATGCGCAGCAGCATG 169
Db      212 CATTAAACCCGTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGTATGCGCAGCAGCATG 271
Qy      170 GTCCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGCAACGCGCATCTGC 229
Db      272 GTCCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGCAACGCGCATCTGC 331
Qy      230 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270
Db      332 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 372
```



```
RESULT 5
US-09-972-916A-2
; Sequence 2, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
US-09-972-916A-2

Query Match      81.1%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 5.8e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TCACAAGCAAAACAACTATTATTGTAACAGCGGGGATCTTAGACGCGCTGACAGATCA 111
DB 1 TCACAGCAAAACAACTATTATTGTAACAGCGGGATCTTAGACGCGCTGACAGATCA 60
QY 112 TTAACCGGTGCTGCGGAGCGAGCCCTTCATMAAGCCCTGGATGAGCCAGCAGCATGCT 171
DB 61 TTAACCGGTGCTGCGGAGCGAGCCCTTCATMAAGCCCTGGATGAGCCAGCAGCATGCT 120
QY 172 CCACGCGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGCC 231
DB 121 CCACGCGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGCC 180
QY 232 AGAGAGCTGTGACGACCACTTCCGCTACTAGCTAGCGCG 270
DB 181 AGAGAGCTGTGACGACCACTTCCGCTACTAGCTAGCGCG 219

RESULT 6
US-09-917-800A-1608
; Sequence 1608, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
```

```
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1608
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match      28.3%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCA 247
DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCA 67
QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 7
US-10-388-934-5
; Sequence 5, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boese, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1 way rat)
US-10-388-934-5

Query Match      28.3%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCA 247
DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCA 67
QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 8
US-10-191-803-73
; Sequence 73, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
```

FILE REFERENCE: 44921-5090US
CURRENT APPLICATION NUMBER: US/10/191,803
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: US 60/303,819
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/305,623
PRIOR FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: US 60/369,351
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/377,611
PRIOR FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 1140
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 73
LENGTH: 1500
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_013144
US-10-191-803-73

Query Match 28.3%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAACCCGAGCAGATTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACAC 247
DB 8 CACAACCCGAGCAGATTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACAC 67

QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 9
US-10-152-319A-1613
Sequence 1613, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgins, Brandon
APPLICANT: Casle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1613

LENGTH: 1500
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match 28.3%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAACCCGAGCAGATTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACAC 247
DB 8 CACAACCCGAGCAGATTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACAC 67

QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 10
US-09-880-107-2393

Sequence 2393, Application US/09880107
Patent No. US20020142981A1

GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2393
LENGTH: 6128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
US-09-880-107-2393

Query Match 22.5%; Score 60.8; DB 9; Length 6128;
Best Local Similarity 66.5%; Pred. No. 8.9e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAAGAAAACAACTATTATTGAACAGCGGATCTAGACGCTGCCGAAATCAT 112
DB 473 CACTAGAAAACAACTATTATTGAACACTCAGCTCAGCGTGGCGCTGCCAATCAT 532

QY 113 TAACCC-----GTGCTGCCAGCAGACCTTATATAGAGCGCTGGTATGCCAGC 162
DB 533 TAACCTCTGTGTAAGTGGCGGCGCTGTGCTTTTAAGTGGCGCTGTCTCAGC 592

QY 163 CAGCATGTCACATGCGCCGCGAGACAAACCCAGGAGCATTTGAACACTGC-ACAAG 221
DB 593 GAGCATGGGCGACCGCATCC-----CATCAGGAGACATGTGCGCGCGCGCG 644

QY 222 CCATCTGCCAGAGAGCTGTGACCACTTC 253
DB 645 CCACCTTCCAGAGAGCATGGCCACCGCTCC 676

RESULT 11
US-10-756-149-1484
Sequence 1484, Application US/10756149
Publication No. US20050181375A1
GENERAL INFORMATION:

```

; APPLICANT: Aziz, Natsasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-1484

Query Match      22.5%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 8-9e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACGCTGCCCTTGACATCAT 112
DB 473 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAAGCCCTGGGTATGGCCAGC 162
DB 533 TAACCTCTGCTGTCAGATGGCGCGGCTGTGCTCTTTATAGGTGGCGCTGTGTCAGC 592
QY 163 CAGCATGTGTCCACTGCGCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 593 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 644
QY 222 CCATCTGCCCGAGAGCTGTGACCACTTC 253
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 12
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CLO00786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-126

Query Match      22.5%; Score 60.8; DB 22; Length 9173;
Best Local Similarity 66.5%; Pred. No. 1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACGCTGCCCTTGACATCAT 112
DB 1877 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCATCAT 1936
QY 113 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAAGCCCTGGGTATGGCCAGC 162
DB 1937 TAACCTCTGCTGTCAGATGGCGCGGCTGTGCTCTTTAAGGTGGCGCTGTGTCAGC 1996
QY 163 CAGCATGTGTCCACTGCGCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 1997 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 2048
```

```

QY 222 CCATCTGCCCGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGCACTGGCCACCGCTCC 2080

RESULT 13
US-10-893-315-160
; Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CLO00786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-160

Query Match      22.5%; Score 60.8; DB 22; Length 9174;
Best Local Similarity 66.5%; Pred. No. 1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACGCTGCCCTTGACATCAT 112
DB 1877 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCATCAT 1936
QY 113 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAAGCCCTGGGTATGGCCAGC 162
DB 1937 TAACCTCTGCTGTCAGATGGCGCGGCTGTGCTCTTTAAGGTGGCGCTGTGTCAGC 1996
QY 163 CAGCATGTGTCCACTGCGCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 1997 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 2048
QY 222 CCATCTGCCCGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGCACTGGCCACCGCTCC 2080

RESULT 14
US-09-972-916A-1
; Sequence 1, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-972-916A-1

Query Match      18.9%; Score 51; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCCGCAGTGTA 51
 |||||
 Db 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCCGCAGTGTA 51

RESULT 15
 US-09-972-916A-4/c
 ; Sequence 4, Application US/09972916A
 ; Patent No. US20020107198A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thule, Peter M.
 ; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
 ; FILE REFERENCE: US 1292/01 (VA)
 ; CURRENT APPLICATION NUMBER: US/09/972,916A
 ; CURRENT FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/239,113
 ; PRIOR FILING DATE: 2000-10-11
 ; NUMBER OF SEQ ID NOS: 6
 ; SEQ ID NO 4
 ; LENGTH: 321
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthesized
 ; US-09-972-916A-4

Query Match 18.9%; Score 51; DB 9; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCCGCAGTGTA 51
 |||||
 Db 102 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCCGCAGTGTA 52

Search completed: September 1, 2005, 16:44:18
 Job time : 793.359 secs

ORGANISM: Human
US-09-949-001-30

TITLE OF INVENTION: WITH OSTEOPOROSIS METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match
Best Local Similarity 66.5%; Score 60.8; DB 4; Length 9174;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGCGGATCTTACACGCTGCCTGACATCAT 112
DB 1877 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCATCAT 1936
QY 113 TAACC-----GTGCTGCCGACGACGCTTCAATAGGCGCTGTGGTATGCGCAGC 162
DB 1937 TAACCTCTGTGCAAGTGGCGGCTGTGCTTATTAAGTGGCGGCTGTGTCAGC 1996
QY 163 CAGCATGTCTCACTGCGCCGCGAGACACAAACCGACGATTTGAACACTGC-ACACGG 221
DB 1997 GAGCATGGCGCACCGGCATC-----CATCGACGAGCATCTGCGCGCGCGCGCG 2048
QY 222 CCATCTGCCAGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGACTGCGCACCGCTCC 2080

RESULT 3
US-08-945-140-1
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Wallstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816

TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match
Best Local Similarity 18.9%; Score 51; DB 3; Length 194;
Matches 51; Conservative 100.0%; Pred. No. 1,4e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGACGGGCACTCCGTTGTCCTGGACTGCGCCGAGTGA 51
DB 11 CATGGGCGACGGGCACTCCGTTGTTCTGACTGCGCCCGAGTGA 61

RESULT 4
US-08-791-849A-14
Sequence 14, Application US/08791849A
Patent No. 591449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,

```
/
/ LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
/ FEATURE:
/ NAME/KEY: Intron
/ LOCATION: join(3219..3765, 3949..5916, 6009..6151,
/ LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..8297
/ LOCATION: 9480..10162)
/ US-08-791-849A-14

Query Match          18.9%; Score 51; DB 2; Length 13011;
Best Local Similarity 100.0%; Pred.No. 1e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGACGCGGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 51
DB 3021 CATGGGCGACGCGGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 3071

RESULT 5
US-09-949-016-11786
/ Sequence 11786, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11786
/ LENGTH: 15108
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-11786

Query Match          13.6%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred.No. 0.096;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 51
DB 2357 CACGGGCGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 2399

RESULT 6
US-09-949-016-17205
/ Sequence 17205, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17205
/ LENGTH: 15108
/ TYPE: DNA

/ ORGANISM: Human
US-09-949-016-17205

Query Match          13.6%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred.No. 0.096;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 51
DB 2357 CACGGGCGACCTCCGTTGCTGACTGTGAGCCCGCCAGTGA 2399

RESULT 7
US-09-949-016-19194/c
/ Sequence 19194, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 19194
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-19194

Query Match          13.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred.No. 0.075;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 130 CCAGCCCTTCATTAAGCCCTGCGTATGCGCAGCAGATGTCACCTCCCGCAGACA 189
DB 478 CCAGCCCTTCCTGGGCGCCTGTTCTTCATTCAGCTTAGCAGCTCCAGGTTAGGCC 419

QY 190 CAACCCGACGAGCATTTGAACATGACACGCGCCATCTGCCAGAGACTGTGACCACCA 249
DB 418 CCCATCCAGAGACATCATCACCAGATGACAGCCCTGCACACCATCTCATATAGCC 359

QY 250 CTT 252
DB 358 CTT 356

RESULT 8
US-09-949-016-73784/c
/ Sequence 73784, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
```


Db 5230 CTT 5228

RESULT 12
US-09-949-016-13895/c
Sequence 13895, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13895
LENGTH: 7562
TYPE: DNA
ORGANISM: Human
US-09-949-016-13895

Query Match
Best Local Similarity 55.3%; Score 35; DB 4; Length 7562;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 130 CCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTGTCACCTGCCCGCCAGACA 189
DB 5350 CCAGGCTTCCTCGGGCGCTGTCTCTCCATCCAGCTTAGCCACCTGCCAGATTAGGCC 5291
QY 190 CAACGCCAGCAGCATTTGAACACTGCACGCGGCATCTGCCAGAGAGCTGTGACACACA 249
DB 5290 CCATTCAGACACATCCACCCGAGATCAGAGCCCTGCCACCACTTCATCATAGACC 5231
QY 250 CTT 252
DB 5230 CTT 5228

RESULT 13
US-08-304-309-3
Sequence 3, Application US/08304309
Patent No. 5856454
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.

REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 88..3162
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..4447
OTHER INFORMATION: /product= "Pig DPD"
US-08-304-309-3

Query Match
Best Local Similarity 49.7%; Score 32.2; DB 2; Length 4447;
Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 96 GCTGCCCTGACATATTAAACCCGTGCTGCCGAGCCACCTTCATTAAGCCCTGGGTAT 155
DB 2275 GATGCTGTACAGCCACCAACCGCTTCAGGTCTCATGGGATTTAAAGCCGATGCGACG 2334
QY 156 GCGCAGCAGCATGTGTCACCTGCCCGCCAGACCAACCCAGCAGCATTTGAACACTGCG 215
DB 2335 CCTGGCCAGCGGTGGGTCTGGCAAGCGACTACATACGAGAGGTCTTGGCAGCGCC 2394
QY 216 ACACGCCCATCTGCCACAGAGCTGTGACCAACCACTTCCGCTACT 260
DB 2395 ATCAGACCAATTGCTTTGAGAGCTGTGACCAACATTCCTCGTCT 2439

RESULT 14
US-08-991-942-3
Sequence 3, Application US/08991942
Patent No. 6015673
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,942
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 88..3162
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..4447
OTHER INFORMATION: /product= "Pig DPD"
US-08-991-942-3

Query Match 11.9%; Score 32.2; DB 3; Length 4447;
Best Local Similarity 49.7%; Pred. No. 1.8;
Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

OY 96 GCTGCCCTTACAATCTATTACCCGCTGCTCCGAGCCCTTTCATTAAGCCCTGGGTAT 155
DB 2275 GATGGTTTACAGCCACCAACGCTCTCAGGTCCTATGAGTTAAAGCCGATGACAG 2334
OY 156 GCCCAGCCAGATGCTCCAGTCCGCGGAGACACAAACCAGCGCATTTGACACTGC 215
DB 2335 CCTGGCCAGCGGTGGTCTGCAAGCGGACTACATACGAGAGTCTTGGCAGGCC 2394
OY 216 ACACGGCCATCTGCCAGAGAGCTGTGACCACTTCCGCTACT 260
DB 2395 ATCAGACCAATTGCTTGTAGAGCTGTGACCACTTGTCTGCT 2439

RESULT 15
US-09-949-016-11745
Sequence 11745, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11745
LENGTH: 36180
TYPE: DNA
ORGANISM: Human
US-09-949-016-11745

Query Match 11.9%; Score 32; DB 4; Length 36180;
Best Local Similarity 53.1%; Pred. No. 5.5;
Matches 68; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 105 ACAATCATTAACCGCTGCTCCGAGCCGCTTCATTAAGCCCTGGGTATGCCAGCCA 164
DB 2652 ACCCCCTTCCCGCGGGGAAGAAAGAGCTAGGTAGGCCCTGGGTTGGGGCCCTA 2711
OY 165 GCATGTCACATGCCCGCGAGACACAAACCAGAGCATTTGAACACTGCACAGGCCA 224
DB 2712 GAGGGTTCACCTCGAGGCCAAGCATGCGCACTGCCCCCAGGGAGAAATCCCTTTGTTTC 2771
OY 225 TCTGCCCA 232
|| |||||

DB 2772 TCCGCCCA 2779

Search completed: September 1, 2005, 07:07:09
Job time : 87.4384 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 1998.69 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321
Sequence: 1 tacactgggggagcagagcgc.....ttccgctactagctagcgcg 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hhg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214.4	66.8	1181	10	RAT1LGFZ
2	214.4	66.8	185148	2	AC136382
3	188.8	58.8	1163	10	MM1LGF
4	188.8	58.8	192843	10	AL607124
5	98.2	30.6	1569	10	BC013345
6	79.2	24.7	1507	10	AY560836
7	76.4	23.8	1500	6	AX401932
8	76.4	23.8	1500	6	AX827271
9	76.4	23.8	1500	10	RAT1LGFZ
10	76.4	23.8	1500	10	RAT1LGFZ
11	76.4	23.8	1500	10	RAT1LGFZ
12	65.4	20.4	1510	10	RAT1LGFZ
13	64	19.9	3886	9	AY095345
14	63.8	19.9	6128	9	AX409747
15	63.8	19.9	6128	9	HM1GFBP1A
16	63.8	19.9	6128	11	G19994
17	63.8	19.9	6480	9	HM1GFBP1
18	63.8	19.9	9082	9	AY434089
19	63.8	19.9	69887	9	AC091524

C 20	62.2	19.4	141539	9	AC146152	Pan trogl
C 21	62.2	19.4	189932	2	AC146117	Pan trogl
C 22	62.2	19.4	200933	2	AC148834	Pan trogl
C 23	60.8	18.9	448	11	G67139	IGFBP1 X1.1
C 24	55.2	17.2	194	6	A57715	Sequence 1
C 25	55.2	17.2	194	6	AR175909	Sequence
C 26	55.2	17.2	13011	6	E14395	gDNA encod
C 27	55.2	17.2	13011	6	AX827302	Sequence
C 28	53.2	17.2	13011	10	RN1PKG	Sequence
C 29	54.4	16.9	895	9	HM1GFBP1	Human insul
C 30	52	16.2	539	11	G67173	Human insul
C 31	47.4	14.8	2717	10	RATPRL1	Human insul
C 32	47.4	14.8	231241	2	AC097039	Rat pyruvat
C 33	44.4	13.8	1480	10	RAT1LGFZ	Rattus norv
C 34	40.4	12.6	415	9	HSPYKINP	Rattus norv
C 35	40.2	12.5	8409	9	HSPYKINP	Rattus norv
C 36	40.2	12.5	13322	9	AY136591	Human sapi
C 37	40.2	12.5	133525	9	AL713999	Human sapi
C 38	37.8	11.8	7061	6	AX251887	Human DNA
C 39	37.8	11.8	7061	6	AX345899	Sequence
C 40	37.8	11.8	7061	6	AX348680	Sequence
C 41	36.8	11.5	348525	1	BX640428	Sequence
C 42	36.8	11.5	348997	1	BX640427	Sequence
C 43	36.8	11.5	349442	1	BX640447	Sequence
C 44	36.6	11.4	1678	6	AX713325	Sequence
C 45	36.6	11.4	1678	9	AK054581	Homo sapi

ALIGNMENTS

RESULT 1
LOCUS RAT1LGFZ
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION M84484
VERSION M84484.1 GI:204927
KEYWORDS insulin-like growth factor.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 1181)
Unceman,T.G., Lacsos,R.G., McGary,E., Whalen,C. and Goswami,R.G.
Biochem. Biophys. Res. Commun. (1991) In press
COMMENT
Original source text: Rattus norvegicus (strain Sprague-Dawley)
male adult liver DNA.

FEATURES
SOURCE
1..1181
Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

Query Match 66.8%; Score 214.4; DB 10; Length 1181;
Best local similarity 97.3%; Pred. No. 4.6e-53;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
ORIGIN
93 TGGCCCATGTACAGCAAACTATTATTGAACACGGGATCTTACAGCTGCC 152
DB 813 TGTGGGAGCTCACAGCAAACTATTATTGAACACGGGATCTTACAGCTGCC 872
OY 153 CTGCAATCATTAACCGGTGTCGCGAGCCCTTCATTAAGCCCTGGTATAGCCAG 212
DB 873 CTGCAATCATTAACCGGTGTCGCGAGCCCTTCATTAAGCCCTGGTATAGCCAG 932
OY 213 CCACATGATGTCACGCGCGGAGACAAACCCAGGAGCATTAAGCATGACACGG 272

Db 993 CCAGCATGTGCTCCAGCTGCCCGCCGAGACACAAACCCAGCAGCATGTGAACACTGCACACG 992
 273 CCATCTGCCAGAGAGCTGTGACCACTTCGGCTACTAGCTA 316
 Db 993 CCATCTGCCAGAGAGCTGTGACCACTTCGGCTACTAGCTA 1036

RESULT 2
 AC136382
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
 AC136382 185148 bp DNA linear HTG 01-NOV-2002
 AC136382
 *** 63 unordered pieces.
 AC136382
 HTG: HTGS PHASE1.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D., Anyalelechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, A., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, U., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kows, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okunolu, G., Olanunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, U., Shvartbeyn, A., Sisson, I., Slater, C. D., Smales, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Sytek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCES
 AUTHORS
 Unpublished
 Rat Genome Sequencing Consortium.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KOGS
 Center clone name: CH230-97018
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap, version 0.990129
 Consensus quality: 135613 bases at least Q40
 Consensus quality: 140849 bases at least Q30
 Consensus quality: 145680 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length.
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a working draft sequence. It currently
 * consists of 63 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1
 1011 1010: contig of 1010 bp in length
 1111 1110: gap of unknown length
 2609 2608: contig of 1498 bp in length
 2709 2708: gap of unknown length
 3916 3915: contig of 1207 bp in length
 4015 4014: gap of unknown length
 5547 5546: contig of 1532 bp in length
 5548 5547: gap of unknown length
 7119 7118: contig of 1472 bp in length
 7219 7218: gap of unknown length
 8885 8884: contig of 1666 bp in length
 8886 8885: gap of unknown length
 10258 10257: contig of 1273 bp in length
 10358 10357: gap of unknown length
 11626 11625: contig of 1268 bp in length
 11627 11626: gap of unknown length
 11727 11726: gap of 1762 bp in length
 13489 13488: gap of unknown length
 13588 13587: gap of unknown length
 15123 15122: contig of 1534 bp in length
 15223 15222: gap of unknown length
 15947 15946: contig of 1724 bp in length
 16947 17046: gap of unknown length
 17047 18164: contig of 1118 bp in length
 18165 18264: gap of unknown length
 19678 19677: contig of 1414 bp in length
 19679 19678: gap of unknown length
 19778 19777: gap of unknown length
 21687 21686: contig of 1909 bp in length
 21688 21787: gap of unknown length
 21788 22892: contig of 1105 bp in length
 22893 22992: gap of unknown length
 22993 24336: contig of 1344 bp in length
 24337 24436: gap of unknown length
 24437 26517: contig of 2181 bp in length
 26518 26717: gap of unknown length
 26718 28577: contig of 1860 bp in length
 28578 28677: gap of unknown length
 28678 30823: contig of 2146 bp in length
 30824 30923: gap of unknown length
 30924 32266: contig of 1343 bp in length
 32267 32366: gap of unknown length
 32367 33894: contig of 1528 bp in length
 33895 33994: gap of unknown length
 33995 35373: contig of 1379 bp in length

35374 35473: gap of unknown length
35474 37295: contig of 1822 bp in length
37296 37395: gap of unknown length
37396 38420: contig of 1025 bp in length
38421 38520: gap of unknown length
38521 39974: contig of 1454 bp in length
39975 40074: gap of unknown length
40075 42283: contig of 2209 bp in length
42284 42383: gap of unknown length
42384 43721: contig of 1338 bp in length
43722 43821: gap of unknown length
43822 46538: contig of 2717 bp in length
46539 46639: gap of unknown length
46639 48621: contig of 1993 bp in length
48622 48721: gap of unknown length
48722 50602: contig of 1881 bp in length
50603 50702: gap of unknown length
50703 53841: contig of 3139 bp in length
53842 53941: gap of unknown length
53942 56195: contig of 2254 bp in length
56196 56295: gap of unknown length
56296 58949: contig of 2654 bp in length
58950 59049: gap of unknown length
59050 61032: contig of 1983 bp in length
61033 61132: gap of unknown length
61133 62942: contig of 1810 bp in length
62943 63043: gap of unknown length
63043 66045: contig of 3003 bp in length
66046 66145: gap of unknown length
66146 68197: contig of 2052 bp in length
68198 68297: gap of unknown length
68298 72235: contig of 3938 bp in length
72236 72335: gap of unknown length
72336 75803: contig of 3468 bp in length
75804 75903: gap of unknown length
75904 78355: contig of 2452 bp in length
78356 78455: gap of unknown length
78456 81458: contig of 3003 bp in length
81459 81558: gap of unknown length
81559 84673: contig of 3115 bp in length
84674 84773: gap of unknown length
84774 87739: contig of 2966 bp in length
87740 87839: gap of unknown length
87840 91535: contig of 3656 bp in length
91536 91635: gap of unknown length
91636 95522: contig of 3887 bp in length
95523 95622: gap of unknown length
95623 98730: contig of 3108 bp in length
98731 98830: gap of unknown length
98831 101618: contig of 2768 bp in length
101619 101718: gap of unknown length
101719 105335: contig of 3817 bp in length
105336 105635: gap of unknown length
105636 109393: contig of 3758 bp in length
109394 109493: gap of unknown length
109494 113564: contig of 4071 bp in length
113565 113664: gap of unknown length
113665 117630: contig of 3966 bp in length
117631 117730: gap of unknown length
117731 122502: contig of 4772 bp in length
122503 122602: gap of unknown length
122603 126544: contig of 4042 bp in length
126545 126744: gap of unknown length

Query Match 66.8%; Score 214.4; DB 2; Length 185148;
Best Local Similarity 97.3%; Pred. No. 6.1e-53;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 93 TCGCCCATGTCAACAAGCAAACTTATTTTGAACAGCGGGATCTAGACGCTGCC 152
DB 99243 TGTGGAGGCTCACAAGCAAACTTATTTTGAACAGCGGGATCTAGACGCTGCC 99302
QY 153 CTGACATCATTAACCCGTCGTCGAGCAGCCCTTCAATAGGCCCTGGGATAGCCAG 212

DB 99303 CTGACATCATTAACCCGTCGTCGAGCAGCCCTTCAATAGGCCCTGGGATAGCCAG 99362
QY 213 CCAGCATGTCACACTGCCCGCGAGACAAACCAGAGAGATTGAACACTGACAGCG 272
DB 99363 CCAGCATGTCACACTGCCCGCGAGACAAACCAGAGAGATTGAACACTGACAGCG 99422
QY 273 CCATTCGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
DB 99423 CCATTCGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 99466

RESULT 3

LOCUS 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
X67493
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Lee, J., Greenbaum, L., Haber, B. A., Nagle, D., Lee, V., Miles, V.,
Mohn, K. L., Bucan, M., and Taub, R.
Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
Hepatology 19 (3), 656-665 (1994)

TITLE
JOURNAL MEDLINE
PUBMED 94164648
7509771

REFERENCE
AUTHORS Mohn, K. L., Waddell, J. R., and Taub, R.
Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries

JOURNAL Nucleic Acids Res.
REFERENCE 3 (bases 1 to 1363)
AUTHORS Taub, R. A.

TITLE Direct Submission
JOURNAL Submitted (23-JUN-1992) R. A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curtis Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source
1. 1363
Location/Qualifiers

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
792..1336
/number=1
792
/note="mRNA cap site"
964..>1336

CDS

/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

/translation="MPEFLTVVSWPFLILLFQIGVAGAPQPMHCAPCTERLGLCP
PVPASCPETSRPAGCCCTCALPMGAAAGVATARCACQGLSRAIPGEPRPLHALTRG
QGASLPEPAPARTSLTFSQHE"

ORIGIN

Query Match 58.8%; Score 188.8; DB 10; Length 1363;
Best Local Similarity 90.2%; Pred. No. 2.1e-45;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 93 TGGCCCATGTGACAGCAAACTATTTTGAACACGGGATCCTAGACGCTGCC 152
|||
DB 667 TGTGTAGAGCTCACAGCAAACTATTTTGAACACGCGGCTCTAGACGCTGCC 726
|||
QY 153 CTGACATCATTTAACCCCTGTGCGGACGACCCCTTCATTAAGGCTGGTATGGCCAG 212
|||
DB 727 CTGACATCATTTAACCTGTGCGGACGACGACCCCTTCATTAAGGCTGGTATGACACG 786
|||
QY 213 CCAGCATGTGTCACTGCGCCGCGAGACACAAACCCAGGACATTGAACACTGCACACG 272
|||
DB 787 CCAGCATGTGTCACTGCGCCGCGAGACACACCCAGGACATTGAACACTGCACACG 846
|||
QY 273 CCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||
DB 847 CCGTCTGCCAGAGAGCTGTGACCACTTCGCTACTACTA 890
|||

RESULT 4
AL607124
LOCUS Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
DEFINITION
ACCESSION AL607124
VERSION AL607124.15 GI:20145926
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Oliver, K.
JOURNAL Direct Submission
Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.choi.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES
source

Location/Qualifiers
1..192843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

/clone="RP23-20C9"
/clone_1ib="RPI-23"

ORIGIN

Query Match 58.8%; Score 188.8; DB 10; Length 192843;
Best Local Similarity 90.2%; Pred. No. 2.8e-45;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 93 TGGCCCATGTGACAGCAAACTATTTTGAACACGGGATCCTAGACGCTGCC 152
|||
DB 2419 TGTGTAGAGCTCACAGCAAACTATTTTGAACACGCGGCTCTAGACGCTGCC 2478
|||
QY 153 CTGACATCATTTAACCCCTGTGCGGACGACCCCTTCATTAAGGCTGGTATGGCCAG 212
|||
DB 2479 CTGACATCATTTAACCTGTGCGGACGACGACCCCTTCATTAAGGCTGGTATGACACG 2538
|||
QY 213 CCAGCATGTGTCACTGCGCCGCGAGACACAAACCCAGGACATTGAACACTGCACACG 272
|||
DB 2539 CCAGCATGTGTCACTGCGCCGCGAGACACACCCAGGACATTGAACACTGCACACG 2598
|||
QY 273 CCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||
DB 2599 CCGTCTGCCAGAGAGCTGTGACCACTTCGCTACTACTA 2642
|||

RESULT 5
BC013345
LOCUS Mus musculus insulin-like growth factor binding protein 1, mRNA
DEFINITION (CDNA clone MGC:14075 IMAGE:4161889), complete cds.
ACCESSION BC013345
VERSION BC013345.1 GI:15426482
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heist, F.,
Dietchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheerz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullan, S.V., Bosak, S.A., McEwen, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huik, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Heltan, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bonfard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1569)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
Submitted (31-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)


```

Query Match          23.8%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 298
      |||||||
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67
      |||||||

QY 299 CACTTCGGCTACTAGCTA 316
      |||||||
DB 68 CACTTCGGCTACTATCTA 85
      |||||||

RESULT 8
AX827271 1500 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 5 from Patent EP1344834.
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
          F. HOFMANN-1A ROCHE AG (CH)
FEATURES
          source
            1. 1500
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          23.8%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 298
      |||||||
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67
      |||||||

QY 299 CACTTCGGCTACTAGCTA 316
      |||||||
DB 68 CACTTCGGCTACTATCTA 85
      |||||||

RESULT 9
RATGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE 91141487
PUBMED 1705004
COMMENT Original source text: Rat, cDNA to mRNA.
FEATURES
          source
            1. 1500
              /organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AA41380.1"
/db_xref="GI:204732"
/translation="MPEFLTVSNPFLILSFQYRVVAGAPQPMHCAPCTAREILCP
PVPASCEISRDPAGCGCCPTCPPLGAAAGVATRCAGLSICALPEPRPLALTRG
QGACVLEPAAVPATSLSGSHEFAKAAVASDELAESBEMTEBOLDSFHMARSRED
QPLIMNAISTYSSMRAREITDCLKKKEPCORELYKLEBRLAAAOOKAGDELYKFLPN
CNKGPFYSKOCETSLDGBAGLCWCYVPMGSKTIPSLFETRGDNCQYFNVQ"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match          23.8%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 298
      |||||||
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67
      |||||||

QY 299 CACTTCGGCTACTAGCTA 316
      |||||||
DB 68 CACTTCGGCTACTATCTA 85
      |||||||

RESULT 10
AX163782 5001 bp DNA linear PAT 22-JUN-2001
LOCUS AX163782
DEFINITION Sequence 46 from Patent WO0138579.
ACCESSION AX163782
VERSION AX163782.1 GI:14544878
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1
AUTHORS Gould-Rothberg, B.E., DiPippo, V.A., Ramegh, T.M. and Gerwein, R.W.
TITLE Method of identifying toxic agents using msal-d-induced differential
JOURNAL Curagen Corporation (US)
FEATURES
          source
            1. 5001
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          23.8%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 9.5e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 298
      |||||||
DB 1 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 60
      |||||||

QY 299 CACTTCGGCTACTAGCTA 316
      |||||||
DB 61 CACTTCGGCTACTATCTA 78
      |||||||

RESULT 11
RATGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE 91141487
PUBMED 1705004
COMMENT Original source text: Rat, cDNA to mRNA.
FEATURES
          source
            1. 1500
              /organism="Rattus norvegicus"

```


Db	1	CACAAACCCAGCGAGCACTTTGAACACTGCACACGGCCACTCTGCCCAGAGAGCTGTGACCAC	60
Qy	299	CACCTTCGGCTACTAGCTA	316
Db	61	CACCTTCGGCTACTATCTA	78
RESULT_12			
LOCUS	BC078889		
DEFINITION	BC078889	1510 bp	mRNA
	Rattus norvegicus insulin-like growth factor binding protein 1,		linear
	mRNA (CDNA clone MGC:93595 IMAGE:7129185), complete cds.		ROD 03-AUG-2004
ACCESSION	BC078889		
VERSION	BC078889.1	GI:50927646	
KEYWORDS	MGC.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
DIFFERENCE	1 (bases 1 to 1510)		

AUTHORS	Kiransberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Sherman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Schaefer, T.E., Brownstein, M.J., Ussid, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McGowan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettlem, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonifard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)
REFERENCE	12477932
AUTHORS	2 (bases 1 to 1510)
TITLE	Director MGC Project.
JOURNAL	Direct Submission
COMMENT	Submitted (02-VUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov
REMARK	Tissue Procurement: Howard Jacobs
COMMENT	cDNA Library Preparation: Express Genomics
REMARK	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
COMMENT	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
REMARK	Web site: http://www-shgc.stanford.edu
COMMENT	Contact: (Dickson, Mark) mcd@paxil.stanford.edu
REMARK	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Series: IRK Plate: 184 Row: e Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6961079.
FEATURES	Location/Qualifiers
source	1. 1510 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116"

gene	/clone="MGC:93595 IMAGE:7129185" /tissue_type="Kidney, rat (Brown Norway)" /clone_id="NH MGC_236" /lab_host="DH10B" /note="Vector: pexpress1" 1..1510 /gene="igfbp1" /note="synonyms: IGFBP, IGF-BP25, IBP1" /db_xref="LocusID:25685" /db_xref="RATMAP:44422" /db_xref="RGD:2872" 142..960 /gene="igfbp1" /codon_start=1 /product="insulin-like growth factor binding protein 1" /protein_id="AAH78889.1" /db_xref="GI:50927647" /db_xref="LocusID:25685" /db_xref="RATMAP:44422" /db_xref="RGD:2872" /translation="MPEFLTVVSWPEFLILSFQVRVYAGAPQPMHCAPCTAERLELCP VPVPSCEPISRPACGCCPTCALPLGAACGVATRCAGLSRCALPGEPRPLHATRG QACVLBPAPATSLSGSHBAKAAVAASDLEASPEMTEBQLSLPHLMPSRED PLIIMAISTYSMSRAREITDILKMKRKCCEYLKULERLAANOQRKSDETVEXYLEN CNKGFPHSKCETSLDGEALCWCVTIPWSKRIPLSLTRGDENCHQYENVON"
CDS	
ORIGIN	
Query Match	20.4%; Score 65.4; DB 10; Length 1510;
Best Local Similarity	98.5%; Pred. No. 1.7e-08;
Matches	66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy	250 CGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGGCTA 309 1 CGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGGCTA 60
Db	310 CTAGCTA 316 61 CTATCTA 67
Oy	310 CTAGCTA 316
Db	61 CTATCTA 67
RESULT 13	
AY095345	
LOCUS	AY095345 3886 bp DNA linear PRI 30-DEC-2002
DEFINITION	Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
ACCESSION	AY095345
VERSION	AY095345.1 GI:20853764
KEYWORDS	Papio anubis (olive baboon)
SOURCE	Papio anubis
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheinae; Papio. 1 (bases 1 to 3886) Kim,J.-J., Taylor,H.-S., Akbas,G.E., Foucher,I., Trembleau,A., Jaffe,R.C., Fazleabas,A.T. and Utemman,T.G. Regulation of insulin-like growth factor binding protein-1 promoter activity by FKHR and HOXA10 in primate endometrial cells Biol. Reprod. 68 (1), 24-30 (2003) 12493691 2 (bases 1 to 3886) Kim,J.-J., Jaffe,R.C. and Fazleabas,A.T. Direct Submission Submitted (11-APR-2002) Department of Obstetrics and Gynecology, University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612, USA
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	Location/Qualifiers 1..3886 /organism="Papio anubis" /mol_type="genomic DNA" /db_xref="taxon:9555" 3655..3886 /gene="IGFBP-1"
gene	

```

mRNA
<3655..>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655..>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAW23273.1"
/db_xref="GI:20853765"
/translation="MSEVPPARVWLVLILLTVQVVTASAPWQAPCSAEKIALCPVPV
PASCEVTRASGCCCPKCALPLGACGVATAR"

ORIGIN
Query Match 19.9%; Score 64; DB 9; Length 3886;
Best Local Similarity 60.6%; Pred. No. 4.8e-08;
Matches 149; Conservative 0; Mismatches 80; Indels 17; Gaps 2;

CDS
69 TCACGAAACACGAGAGTGCCTGGCCGATGTCAACAGCAAAACAACTTATTTTGA 128
Db TCCTCCACCTCGCGGTTTGGCTGTAAGGCTTGGCGCACTACAAAACAACTTATTTTGA 339
129 ACACGGGGATCCTAGACGCTGCGCCCTGACATCATTAACCC-----GTGCTGCCG 178
Db ACACCTGAGCTCTTAGCGCGCGCGCTGCGCATCTTAACCTCTGTCGCAAGTGGCGCG 345
179 AGCAGCCCTTCATPAAGGCTCTGGGTATAGCGCAGCCAGCATGTCCACTGCCGCCGAGA 238
Db CCTGGCCCTTATTAAGGCGCGCGCTGTGTCCAGCAAGCATGGCCACCGCATCTCCATC 3510
239 CACAAACCCAGCAGCATTTGAACACTGCACACGCGCCATCTGCCAGAGCTGACAC 298
Db CAGCAAG-----ATCTGCCGCGCGCGCGCCGACCTTCCAGAGACACTGGCCAC 3563
299 CACTTC 304
Db CGCTCC 3569

RESULT 14
LOCUS AX409747 6128 bp DNA linear PAT 14-UN-2002
DEFINITION Sequence 2394 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:21442452
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Alvarez,C., Horne,D., Peters-da-Silva,S. and Vockley,J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2394 11-APR-2002;
JOURNAL GENE LOGIC INC (US)
FEATURES
source
1..6128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/Genbank Accession No. W74587"

ORIGIN
Query Match 19.9%; Score 63.8; DB 6; Length 6128;
Best Local Similarity 63.2%; Pred. No. 5.6e-08;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
69 TCACGAAACACGAGAGTGCCTGGCCGATGTCAACAGCAAAACAACTTATTTTGA 128
Db TCCTCCACCTCGCGGTTTGGCTGTAAGGCTTGGCGCACTACAAAACAACTTATTTTGA 497
438 TCCTCCACCAACGCGGTTTGGCTGTAAGGCTTGGCGCACTACAAAACAACTTATTTTGA 497
129 ACACGGGGATCTTAGACGCTGCCCTTGAACAATTAATTAACCC-----GTGCTGCCG 178
Db ACACCTGAGCTCTTAGCGCGCGCGCTGCGCATCTTAACCTCTGTCGCAAGTGGCGCG 557
498 ACACCTGAGCTCTTAGCGCGCGCGCTGCGCATCTTAACCTCTGTCGCAAGTGGCGCG 557

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 2440.45 Seconds
(without alignment)
5006.706 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321

Sequence: 1 taccactggggccagagatcc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hlc: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	30.6	1013	2	BF236974
2	87.8	27.4	310	6	BY783538
3	87.8	27.4	331	6	BY794229
4	87.8	27.4	354	6	BY771317
5	86.8	27.0	401	1	AI785818
6	86.8	27.0	480	1	AI196314
7	86.8	27.0	706	1	AI530146
8	86.8	27.0	765	1	AI098594
9	86.8	27.0	785	1	AI530313
10	86.8	27.0	799	1	AI529939
11	86.8	27.0	811	6	CA478518
12	86.8	27.0	848	1	AI790802
13	86.8	27.0	852	1	AI528304
14	85.2	26.5	305	1	AI196154
15	82.2	25.6	380	1	AI785039
16	81.4	25.4	846	7	COS73026
17	80.4	25.0	692	7	CV127049
18	79.4	24.7	713	7	COS60662
19	75.4	23.5	748	7	COS75629
20	73.8	23.0	694	7	CV117001
21	72.4	22.8	632	7	W30013
22	72.4	22.6	615	7	BB660958
23	72.4	22.6	618	6	CD561711
24	72.4	22.6	632	1	AI892189

25	67.4	21.0	488	1	AA674302
26	65.4	20.4	801	7	CK472246
27	64.4	20.1	840	7	CK473709
28	62.8	19.6	759	1	AA105355
29	61.8	19.3	269	2	BB604790
30	61.6	19.2	587	2	AW916227
31	60.8	18.9	1589	3	CR621807
32	60.8	18.9	1601	3	CR595377
33	59	18.4	545	1	AA060360
34	47.2	14.7	659	8	AZ840793
35	38.8	12.1	698	5	BU684425
36	36.8	11.5	428	2	BF829782
37	36.8	11.5	567	2	CV014020
38	36.8	11.5	849	9	CNS0206E
39	36.6	11.4	370	7	CN402791
40	36.6	11.4	459	7	CN402791
41	36.6	11.4	466	7	CN402792
42	36.6	11.4	497	6	CD635411
43	36.6	11.4	498	6	CD635412
44	36.6	11.4	744	7	CN402794
45	36.6	11.4	771	5	BX360059

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:416189 5',
LOCUS BF236974
mRNA sequence.

ACCESSION BF236974.1 GI:11150891
VERSION BF236974.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence at: 581.

FEATURES

source location/Qualifiers
1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:416189"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6, site 1: NotI; site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 30.6%, Score 98.2, DB 2, Length 1013;
Best Local Similarity 92.8%, Pred. No. 2.6e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 206 TGGCAGGACAGATGTCCTGCGCCGCGAGACACAAACCGAGACATTGAACACTG 265

Db	9	TGGGAGCCAGCATGTGTCACCTGCCCCGGGAGACACACACCCAGCGAGATTGACACTG	68
0y	266	CACACGGCCATCTGCCAGAGAGCTGTGACCAACCACTTCGGTACTAGCTA	316
Db	69	CACACGGCCGCTGTGCCAGAGAGCTGTGACCAACCACTTCGGTACTAGCTA	119
RESULT 2			
BY783538			
LOCUS		310 bp	mRNA
DEFINITION	BY783538	RIKEN full-length enriched, 17.5 days embryo whole body	EST 23-MAR-2004
ACCESSION	Mus musculus	CDNA clone U930176D05 5', mRNA sequence.	
KEYWORDS	BY783538		
SOURCE	BY783538.1	GI:39710177	
ORGANISM	Mus musculus	(house mouse)	
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 310)		
	Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Otsu, N., Fukuda, S., Sato, K., Watanabe, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Shigenaga, S., Beisel, K., Pavan, W., Aldrich, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakamura, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, Y., Mommaerts, P., Muramatsu, M., Okazaki, Y., Kawai, U. and Hayashizaki, Y.		
TITLE	Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res.	13 (6B), 1273-1289	(2003)
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shuhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details. Location/Qualifiers 1. 310		
FEATURES			
source	/organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="U930176D05" /tissue_type="whole body" /dev_stage="17.5 days embryo" /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	27.4%	Score 87.8	DB 6
Best Local Similarity	92.9%	Pred. No. 2,1e-14	
Matches	92	Conservative 0	Mismatches 7
		Indels 0	Gaps 0
0y	218	ATGGTCACCTGCGCGCGAGACACAAACCGCGAGCTTGAACACTGACACGGCCATC	277
Db	2	ATGGTCACCTGCGCGCGAGACACACCCACCGAGCATTTGAACACTGACACGGCCGCT	61
0y	278	TGCCAGAGAGCTGTGACACCACTTCGGCTACTAGCTA	316

Db	62	TGCCAGAGAGCTGTGACCACCATTCGCCACTACTATCTTA	100
RESULT 3			
BY944229			
LOCUS			
DEFINITION	BY944229 RIKEN full-length enriched, 17.5 days embryo whole body	331 bp	mRNA linear EST 23-MAR-2004
ACCESSION	BY944229		
VERSION	BY944229.1	GI:39720868	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia, Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 331) Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Alizawa, K., Arai, K., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanabe, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Payan, W., Aidinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawaji, J. and Hayashizaki, Y.		
TITLE	Targeting a complex transposome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Saitoh-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES			
source	Location/Qualifiers		
	1..331		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strat="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="L930292H24"		
	/tissue type="whole body"		
	/dev stage="17.5 days embryo"		
	/clone.lib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	27.4%	Score 87.8;	DB 6; Length 331;
Best Local Similarity	92.9%	Pred. No. 2,1e-14;	
Matches	92; Conservative	0; Mismatches	7; Indels 0; Gaps 0;
QY	218	ATGTGTCACGTGCGCCGAGACACAAACCCAGCGACGACATTGAAACACTGACACAGGCCATC	277
DB	2	ATGTGTCACGTGCGCCGAGAGACACACACCCAGCAGACATGGAACACTGACACAGCGCGTC	61
QY	278	TGCCAGAGAGCTGTGACACCACTTCGCGACTACAGCTA	316
DB	62	TGCCAGAGAGCTGTGACACCACTTCGCGACTACTATCTTA	100
RESULT 4			
BY771317			

LOCUS BY7171317 354 bp mRNA linear EST 23-MAR-2004
 DEFINITION BY7171317 RIKEN full-length enriched, 17.5 days embryo whole body
 ACCESSION Mus musculus cDNA clone U930052G15 5', mRNA sequence.
 VERSION BY7171317
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 354)
 REFERENCE Carinini, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugihara, Y., Saito, R., Oosato, N., Fukuda, S., Sato, K., Matsubara, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kuehabe, M., Guetschich, S., Beisel, K., Pavan, W., Aldinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakaguchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
 Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia
 Genome Res. 13 (6B), 1273-1289 (2003)
 JOURNAL MEDLINE 22703353
 PUBMED 12819125
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-research.riken.jp, URL: http://genome.gsc.riken.jp/
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.jp/) for further details.
 FEATURES
 source Location/Qualifiers
 1..354
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U930052G15"
 /tissue_type="whole body"
 /dev_stage="17.5 days embryo"
 /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"
 ORIGIN
 Query Match 27.4%; Score 87.8; DB 6; Length 354;
 Best Local Similarity 92.9%; Pred. No. 2.1e-14;
 Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 218 ATGTTCACCTGCCCGGAGACACAAACCCAGCAGCATTTGAACCTGCACACGGCCATC 277
 |||||
 DB 2 ATGTTCACCTGCCCGGAGACACAAACCCAGCAGCATTTGAACCTGCACACGGCCATC 277
 |||||
 QY 278 TGCCGAGAGAGCTGTGACCAACCACTTCGCTCTACTAGCTA 316
 |||||
 DB 62 TGCCGAGAGAGCTGTGACCAACCACTTCGCTCTACTAGCTA 100
 |||||
 RESULT 5
 LOCUS A1785818 401 bp mRNA linear EST 02-JUL-1999
 DEFINITION u1785818.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1888569 5', similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1785818

VERSION A1785818.1 GI:533534
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 401)
 REFERENCE Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritten, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 JOURNAL Other ESTs: u1785818.x1
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@washington.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972893
 Seq primer: custom primer used
 High quality sequence stop: 126.
 FEATURES
 source Location/Qualifiers
 1..401
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1888569"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-PL3; Site: 1. DraIII (CACTGTG); Site: 2. DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 (ATGTGCGCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCCTACTGG) digested and cloned into distinct DraIII sites of the pME18S-PL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCGCTCTAAAGCTCG and 3' end primer CGACCTCGACCTGACGACA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 401;
 Best Local Similarity 92.9%; Pred. No. 4.2e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCCCGGAGACACAAACCCAGCAGCATTTGAACCTGCACACGGCCATC 278
 |||||
 DB 1 TGTTCACCTGCCCGGAGACACAAACCCAGCAGCATTTGAACCTGCACACGGCCATC 278
 |||||
 QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGCTCTACTAGCTA 316
 |||||
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTCTACTAGCTA 98
 |||||
 RESULT 6
 LOCUS A1196314 480 bp mRNA linear EST 14-OCT-1998
 DEFINITION u1196314.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1887828 5', similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1196314

ACCESSION A1196314
 VERSION A1196314.1 GI:3748920
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, U., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972152
 Seg primer: custom primer used
 High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGCTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTCTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGCTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACTCGACTCGAGCA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 4.4e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCCCCCGAGACACAACCCAGGACATTGAACATCTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCCCCCGAGACACACACCCAGGACATTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACACACCACTTCGGCTACTAGCTA 316
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGGCTACTACTTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION U99F09.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:M5931c_rna1 INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146
 VERSION A1530146.1 GI:4444281
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 Seg primer: custom primer used
 High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGCTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTCTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGCTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACTCGACTCGAGCA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 4.7e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCCCCCGAGACACAACCCAGGACATTGAACACTCTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCCCCCGAGACACACACCCAGGACATTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACACACCACTTCGGCTACTAGCTA 316
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGGCTACTACTTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION U61E07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579 M. musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 765)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930344
 Seq primer: custom primer used
 High quality sequence stop: 395.
 FEATURES
 source
 1..765
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1481988"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18-F13; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 (ATGTGGCCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCACTGG) digested and cloned into distinct DraIII sites of the pME18-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 765;
 Best Local Similarity 92.9%; Pred. No. 4.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCGCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCGCCGCGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 316
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 98
 RESULT 9
 A1530313 785 bp mRNA linear EST 18-MAR-1999
 LOCUS
 DEFINITION u187c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA
 sequence.

ACCESION A1530313
 VERSION A1530313.1 GI:4444448
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 785)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence stop: 459.
 FEATURES
 source
 1..785
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889785"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18-F13; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 (ATGTGGCCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCCACTGG) digested and cloned into distinct DraIII sites of the pME18-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 785;
 Best Local Similarity 92.9%; Pred. No. 4.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCGCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCGCCGCGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 316
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 98
 RESULT 10
 A1529939 799 bp mRNA linear EST 18-MAR-1999
 LOCUS
 DEFINITION u187c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA
 sequence.

ACCESSION AI529939
 VERSION AI529939.1 GI:4444074
 KEYWORDS EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 1 (baes 1 to 799)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
 Person, B., Smaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R.,
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra W/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 Location/Qualifiers
 1..799

JOURNAL
COMMENT

Unpublished (1996)
Contact: Maira M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:971987

FEATURES

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers
1..605

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:187663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTT] (TTTTTTT); double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGACA."

```

ORIGIN

Query Match

Best Local Similarity 26.5%; Score 85.2; DB 1; Length 605;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 219 TGGTCACATCCCGCGAGACACAAACCGAGCATTTGAACACTGCACAGGCGCATCT 278
|||||
DB 1 TGGTCACATCCCGCGAGACACACACCCAGGAGATTGAACACTGCACAGGCGCATCT 60
|||||
QY 279 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98
|||||

RESULT 15

AT785039

LOCUS 380 bp mRNA linear EST 02-JUL-1999
DEFINITION ui73a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888018.5' similar to gb:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION

AT785039
AT785039.1 GI:5332755

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Mus musculus (house mouse)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Maira, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterson, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

COMMENT

Other ESTs: ui73a06.x1
Contact: Maira M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:972342

FEATURES

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers
1..380

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTT] (TTTTTTT); double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGACA."

```

ORIGIN

Query Match

Best Local Similarity 25.6%; Score 82.2; DB 1; Length 380;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 222 TCCACTGCCCGCGAGACACAAACCGAGCATTTGAACACTGCACAGGCGCATTTGCC 281
|||||
DB 1 TCCACTGCCCGCGAGAGACACACACCCAGGAGATTGAACACTGCATTCGCGCTGCC 60
|||||
QY 282 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||||
DB 61 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 95
|||||

Search completed: September 1, 2005, 06:56:49
odb time : 2441.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 934.693 Seconds
(without alignments) 2248.600 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321

Sequence: 1 tacactggggggcagagtcgc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 327445616 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321	100.0	321	9	US-09-972-916A-4
2	321	100.0	372	9	US-09-972-916A-5
3	219.4	66.3	270	9	US-09-972-916A-3
4	219.4	66.3	423	9	US-09-972-916A-6
5	219	66.2	219	9	US-09-972-916A-2
6	103	33.1	423	9	US-09-972-916A-6
7	76.4	23.8	1500	9	US-09-917-800A-1608

8	76.4	23.8	1500	17	US-10-388-934-5	Sequence 5, Appli
9	76.4	23.8	1500	17	US-10-191-803-73	Sequence 73, Appl
10	76.4	23.8	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
11	63.8	19.9	6128	9	US-09-880-107-2393	Sequence 2393, Ap
12	63.8	19.9	6128	22	US-10-756-149-1484	Sequence 1484, Ap
13	63.8	19.9	9173	22	US-10-893-315-126	Sequence 126, App
14	63.8	19.9	9174	22	US-10-893-315-160	Sequence 160, App
15	55.2	17.2	13011	17	US-10-388-934-36	Sequence 36, Appl
16	55.2	17.2	13011	18	US-10-152-319A-2130	Sequence 2150, Ap
17	51	15.9	51	9	US-09-972-916A-1	Sequence 1, Appl1
18	51	15.9	270	9	US-09-972-916A-3	Sequence 3, Appl1
19	40.2	12.5	31703	17	US-10-085-117-172	Sequence 172, App
20	37.8	11.8	7061	15	US-10-311-455-970	Sequence 970, App
21	37.8	11.8	7061	17	US-10-221-613-148	Sequence 148, App
22	36.6	11.4	1678	17	US-10-094-749-9	Sequence 9, Appl1
23	36	11.2	761	17	US-10-264-237-828	Sequence 828, App
24	35.6	11.1	1490	20	US-09-918-995-22070	Sequence 22070, A
25	35.2	11.0	1687	20	US-10-739-930-5461	Sequence 5461, Ap
26	35	10.9	769	13	US-10-027-632-164336	Sequence 164336,
27	35	10.9	769	17	US-10-027-632-164336	Sequence 164336,
28	35	10.9	3895	14	US-10-011-585A-76	Sequence 14335,
29	34	10.6	49979	19	US-10-741-601-5746	Sequence 746, Appl
30	34	10.6	49979	21	US-10-741-600-17905	Sequence 17905, A
31	32.8	10.2	2130	9	US-09-909-962A-7	Sequence 7, Appl1
32	32.8	10.2	2130	9	US-10-437-863-79182	Sequence 79182, A
33	32.4	10.1	1389	19	US-10-437-863-79182	Sequence 79185, A
34	32.4	10.1	1584	19	US-10-437-863-79185	Sequence 1, Appl1
35	32.4	10.1	30000	11	US-09-980-217-1	Sequence 25562, A
36	32.2	10.0	750	13	US-10-027-632-25562	Sequence 25563, A
37	32.2	10.0	750	13	US-10-027-632-25562	Sequence 25564, A
38	32.2	10.0	750	13	US-10-027-632-25564	Sequence 25564, A
39	32.2	10.0	750	17	US-10-027-632-25562	Sequence 25563, A
40	32.2	10.0	750	17	US-10-027-632-25563	Sequence 25563, A
41	32.2	10.0	750	17	US-10-027-632-25563	Sequence 25563, A
42	32.2	10.0	755	13	US-10-027-632-150407	Sequence 150407,
43	32.2	10.0	755	17	US-10-027-632-150407	Sequence 150407,
44	32	10.0	76180	19	US-10-322-881-492	Sequence 492, App
45	32	10.0	91760	13	US-10-087-192-844	Sequence 844, App

ALIGNMENTS

RESULT 1
US-09-972-916A-4
Sequence 4, Application US/09972916A
Patent No. US20020107198A1
GENERAL INFORMATION:
APPLICANT: Thule, Peter M.
TITLE OF INVENTION: GUCCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
FILE REFERENCE: US 1292/01 (VA)
CURRENT APPLICATION NUMBER: US/09/972, 916A
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/239, 113
PRIOR FILING DATE: 2000-10-11
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 4
LENGTH: 321
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 100.0%; Score 321; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 9, 2e-100; Indels 0; Gaps 0;
Matches 321; Conservative 0; Mismatches 0;

QY 1 TACACTGGGGGCGCAGAGTCGAGAACCAACGAGAGTCCCGCCCATGTACACTGGG 60
DB 1 TACACTGGGGGCGCAGAGTCGAGAACCAACGAGAGTCCCGCCCATGTACACTGGG 60
QY 61 GGCCAGAGTCGAGAACCAACGAGAGTCCCGCCCATGTACAGCAAGAAACAAACT 120

```
Db      61 GGCACAGATCCAGGAACCAACGGAGATGCCCCGCGCCCATGATGCAACAAACAACT 120
        |||
Qy      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCAATTAACCCGTCCTGCCAG 180
        |||
Db      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCAATTAACCCGTCCTGCCAG 180
        |||
Qy      181 CCAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCCAGCATGTCATCACTGCCCGCAGACA 240
        |||
Db      181 CCAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCCAGCATGTCATCACTGCCCGCAGACA 240
        |||
Qy      241 CAAACCCAGCGAGCATTTGAACACTGCAACGCGCATCTGCCCAAGAGCTGTGACCAACA 300
        |||
Db      241 CAAACCCAGCGAGCATTTGAACACTGCAACGCGCATCTGCCCAAGAGCTGTGACCAACA 300
        |||
Qy      301 CTTCCGCTACTAGTAGCCGC 321
        |||
Db      301 CTTCCGCTACTAGTAGCCGC 321
        |||
```

```
RESULT 2
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5
```

```
Query Match      100.0%; Score 321; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 9.5e-100;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TACACTGGGGGCGCAGAGTCCAGAGTCCAGAGTGCCTGCGCCCATGTAACACTGG 60
        |||
Db      52 TACACTGGGGGCGCAGAGTCCAGAGTCCAGAGTGCCTGCGCCCATGTAACACTGG 111
        |||
Qy      61 GGCAGAGTCCAGGAACCAACGGAGTGCCTGCGCCCATGTAACAAACAAACT 120
        |||
Db      112 GGCAGAGTCCAGGAACCAACGGAGTGCCTGCGCCCATGTAACAAACAAACT 171
        |||
Qy      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCAATTAACCCGTCCTGCCAG 180
        |||
Db      172 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCAATTAACCCGTCCTGCCAG 231
        |||
Qy      181 CCAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCCAGCATGTCATCACTGCCCGCAGACA 240
        |||
Db      232 CCAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCCAGCATGTCATCACTGCCCGCAGACA 291
        |||
Qy      241 CAAACCCAGCGAGCATTTGAACACTGCAACGCGCATCTGCCCAAGAGCTGTGACCAACA 300
        |||
Db      242 CAAACCCAGCGAGCATTTGAACACTGCAACGCGCATCTGCCCAAGAGCTGTGACCAACA 351
        |||
Qy      301 CTTCCGCTACTAGTAGCCGC 321
        |||
Db      352 CTTCCGCTACTAGTAGCCGC 372
        |||
```

```
RESULT 3
US-09-972-916A-3
; Sequence 3, Application US/09972916A
```

```
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3
```

```
Query Match      68.3%; Score 219.4; DB 9; Length 270;
Best Local Similarity 99.5%; Pred. No. 6.7e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      101 TGTCAAGCAAAACAACTTATTTGAACAGCGGAGTCTTAGACGCTGCTGACAAAT 160
        |||
Db      50 TATCAAGCAAAACAACTTATTTGAACAGCGGAGTCTTAGACGCTGCTGACAAAT 109
        |||
Qy      161 CATTAACCCGTCGTCGCGAGCGAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCAGCATG 220
        |||
Db      110 CATTAACCCGTCGTCGCGAGCGAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCAGCATG 169
        |||
Qy      221 GTCCACTGCCCCCGCAGACACAAACCCAGCAGCATTTGAACATGCAACAGCCCATCTGC 280
        |||
Db      170 GTCCACTGCCCCCGCAGACACAAACCCAGCAGCATTTGAACATGCAACAGCCCATCTGC 229
        |||
Qy      281 CCAGAGCTGTGACACCACTTCGCTACTAGTAGCCGC 321
        |||
Db      230 CCAGAGCTGTGACACCACTTCGCTACTAGTAGCCGC 270
        |||
```

```
RESULT 4
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6
```

```
Query Match      68.3%; Score 219.4; DB 9; Length 423;
Best Local Similarity 99.5%; Pred. No. 7.5e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      101 TGTCAAGCAAAACAACTTATTTGAACAGCGGAGTCTTAGACGCTGCTGACAAAT 160
        |||
Db      203 TATCAAGCAAAACAACTTATTTGAACAGCGGAGTCTTAGACGCTGCTGACAAAT 262
        |||
Qy      161 CATTAACCCGTCGTCGCGAGCGAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCAGCATG 220
        |||
Db      263 CATTAACCCGTCGTCGCGAGCGAGCCCTTCATTAAGGCTCTGGGTATGCGCCAGCAGCATG 322
        |||
Qy      221 GTCCACTGCCCCCGCAGACACAAACCCAGCAGCATTTGAACACTGCAACAGCCCATCTGC 280
        |||
```

Db 323 GTCTCATGCCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGACACAGGCCATCTGC 382
Qy 281 CCAGAGAGCTGTGACCAACCACTTCGCTTACTAGTACGCCG 321
Db 383 CCAGAGAGCTGTGACCAACCACTTCGCTTACTAGTACGCCG 423

RESULT 5

US-09-972-916A-2
/ Sequence 2, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 2
/ LENGTH: 219
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
US-09-972-916A-2

Query Match 68.2%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.7e-65;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 TCACAAGCAAAACAATTTTGAACACGGGAGTCTTACAGACGCTGCTTGAACATCA 162
Db 1 TCACAAGCAAAACAATTTTGAACACGGGAGTCTTACAGACGCTGCTTGAACATCA 60
Qy 163 TTAACCCGCTGTCGCGAGCCCTTCATAGGCCCTGGTATGCGCCAGCAGATGCT 222
Db 61 TTAACCCGCTGTCGCGAGCCCTTCATAGGCCCTGGTATGCGCCAGCAGATGCT 120
Qy 223 CCAGTCCCGCGGAGACCAAAACCCAGCAGCATTTGAACACTGACACAGGCCATCTGCC 282
Db 121 CCAGTCCCGCGGAGACCAAAACCCAGCAGCATTTGAACACTGACACAGGCCATCTGCC 180
Qy 283 AGAGAGCTGTGACCAACCACTTCGCTTACTAGTACGCCG 321
Db 181 AGAGAGCTGTGACCAACCACTTCGCTTACTAGTACGCCG 219

RESULT 6

US-09-972-916A-6/c
/ Sequence 6, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 6
/ LENGTH: 423
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match 32.1%; Score 103; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 6.9e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTAGGGGGCCAGAGTTCAGAAACAGGGAGTCCCCGTCGCCCATTTACACTGGG 60
Db 204 TACTAGGGGGCCAGAGTTCAGAAACAGGGAGTCCCCGTCGCCCATTTACACTGGG 145
Qy 61 GGCCAGAGTCCAGAAACAGGGAGTGGCCCGTGGCCCATGT 103
Db 144 GGCCAGAGTCCAGAAACAGGGAGTGGCCCGTGGCCCATGT 102

RESULT 7

US-09-917-800A-1608
/ Sequence 1608, Application US/09917800A
/ Patent No. US20020119462A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Porter, Mark
/ APPLICANT: Johnson, Kory
/ APPLICANT: Castle, Arthur
/ APPLICANT: Elashoff, Michael
/ TITLE OF INVENTION: Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5038-US
/ CURRENT APPLICATION NUMBER: US/09/917,800A
/ PRIOR FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,040
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,880
/ PRIOR FILING DATE: 2000-11-02
/ PRIOR APPLICATION NUMBER: US 60/290,029
/ PRIOR FILING DATE: 2001-05-11
/ PRIOR APPLICATION NUMBER: US 60/290,645
/ PRIOR FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: US 60/292,336
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: US 60/295,798
/ PRIOR FILING DATE: 2001-06-06
/ PRIOR APPLICATION NUMBER: US 60/297,457
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,884
/ PRIOR FILING DATE: 2001-06-19
/ PRIOR APPLICATION NUMBER: US 60/303,459
/ PRIOR FILING DATE: 2001-07-09
/ NUMBER OF SEQ ID NOS: 1740
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1608
/ LENGTH: 1500
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
/ OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match 23.8%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 CACAAACCCAGCAGCATTTGAACACTGACACAGGCCATCTGCCAGAGAGCTGTGACCAG 298
Db 8 CACAAACCCAGCAGCATTTGAACACTGACACAGGCCATCTGCCAGAGAGCTGTGACCAG 67
Qy 299 CACTTCGCTTACTAGCTA 316
Db 68 CACTTCGCTTACTAGCTA 85

RESULT 8

US-10-386-934-5
/ Sequence 5, Application US/10386934
/ Publication No. US20040005547A1
/ GENERAL INFORMATION:
/ APPLICANT: Boess, Franziska
/ APPLICANT: Suter-Dick, Laura

```

; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US2004000547A1way rat)
US-10-388-934-5

Query Match      23.8%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 298
      |||
Db      8 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
      |||

Qy      299 CACTTCGCTACTACTA 316
      |||
Db      68 CACTTCGCTACTACTA 85
      |||

RESULT 9
US-10-191-803-73
; Sequence 73, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_013144
US-10-191-803-73

Query Match      23.8%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 298
      |||
Db      8 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
      |||

Qy      299 CACTTCGCTACTACTA 316
      |||
Db      299 CACTTCGCTACTACTA 316
      |||
```

```

Db      68 CACTTCGCTACTACTA 85
      |||

RESULT 10
US-10-152-319A-1613
; Sequence 1613, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1613
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match      23.8%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 298
      |||
Db      8 CACAACCCGAGCATTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
      |||

Qy      299 CACTTCGCTACTACTA 316
      |||
Db      68 CACTTCGCTACTACTA 85
      |||

RESULT 11
US-09-880-107-2393
; Sequence 2393, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
```



```

1 CURRENT APPLICATION NUMBER: US/09/860,107
2 CURRENT FILING DATE: 2001-06-14
3 PRIOR APPLICATION NUMBER: US 60/211,379
4 PRIOR FILING DATE: 2000-06-14
5 PRIOR APPLICATION NUMBER: US 60/237,054
6 PRIOR FILING DATE: 2000-10-02
7 NUMBER OF SEQ ID NOS: 3950
8 SOFTWARE: PatentIn Ver. 2.1
9 SEQ ID NO 2393
10 LENGTH: 6128
11 TYPE: DNA
12 ORGANISM: Homo sapiens
13 FEATURE:
14 OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
15 US-09-860-107-2393

```

Query Match	19.9%	Score 63.8	DB 9	Length 6128
Best Local Similarity	63.2%	Pred. No. 3	8e-11	
Matches	156	Conservative	0	Mismatches 72; Indels 19; Gaps 3

QY	69	TCACAGAAACA	CGGGAGTGC	CCCCGTGC	CGCCCATGT	CAACAAGAAACAAACTTATTTTGA	128
Db	438	TCCCTCCACCA	CGCGGTTTGC	GTAGGCGCTT	TGGGTGCAC	TACGAAACAAACTTATTTTGA	497
QY	129	ACACGGGGAT	CTTAGCAGCTGC	TGCCCTGA	CAATCATTTAACCC-----GTGGTGGCGG	178	
Db	498	ACACTCAGCT	CTTAGCGGTGGCGG	GTGGCAATCATTTAACCTT	CTGTGTCGAAAGTGGCGCGG	557	
QY	179	AGCCAGCCCTT	CAITAAAGGCCCTT	GGGTATGAGCC	AGCAGCATGTC	CACTGCCCCGCGAGA	238
Db	558	CCTGTGCCCC	TTTATTAAGGTGGCGG	CTGTGTGTCC	AGAGAGATGGGC	ACCGCCATCC----	613
QY	239	CACAAACCCAG	CGAGCATTTGAACA	CTGTC-AC	ACGGSCCACTCTCC	CCAGAGAGCTGTGACCA	297
Db	614	-----CAT	CAAGAGACATCTG	CGCCGCGCGCCG	CCGCCCACTCC	CCAGAGAGCATCTGCGCA	669
QY	298	CCACTTC	304				
Db	670	CCGCTCC	676				

```

RESULT 12
US-10-756-149-1484
; Sequence 1484, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-756-149-1484

```

	Query Match	Best Local Similarity	Score	DB	Length
Matches	156;	Conservative	0;	Mismatches	72;
				Indels	19;
				Gaps	3
QY	69	TTCCAGGAAACACGGGAGTGTCCCGCGGCCCATGTGCACAGCAAAACCTATTATTGGA	128		
Db	438	TCCTCCACACAGCGGTTTGGCTGAGGCGCTTGGGTGCACCTAGCAAAACCTATTATTGGA	497		
QY	129	ACACGGGGATCTTGCACACGCTGCTCGTCAATCATTAACCC-----GTGTGCGG	178		
Db	498	ACACTCTGACTCTTACGCGCGCGCTGTGCATTCATTAACCTCTGTCGCAATGTCGCGG	557		

Qy 173 AGCGAGCCCTTATATAGGCGCTGGGTATAGGCGACGACATGTCTCACTGCCGCGGAGA 238
 |||||
 Db 558 CCTGTGCCCTTATATAGGCGCGCGGTGTGTGCAGCGACATCGCGCACCGCCATTC--- 613
 Qy 239 CACAAACCGAGGAGCATTTGAACACTGC-ACACGCGCATCTGCCCAAGAGCTGTGACCA 297
 |||||
 Db 614 ----CATTCAGGAGCACTCTGCGCGCGCGCGCGCCACCTCTCCAGAGAGACTGTGCCA 669
 Qy 298 CCACTTC 304
 |||||
 Db 670 CCGCTCC 676

```

RESULT 13
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication NO. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
; US-10-893-315-126

```

Query Match	19.9%	Score 63.6	DB 22	Length 9173
Best Local Similarity	63.2%	Pred. No. 4,38-11		
Matches 155	Conservative 0	Mismatches 72	Indels 19	Gaps 3
QY	69	TCACGAGAACCA	CGGGAATGCCCGCTGGCGCCCATCTCACAGCAAAACAACCTATTATTGGA	128
Db	1842	TCTCTCCACACAGGGGTTGGGTAGGGGCTTGGGTGACACTAGCAAAACAACTTATTATTGGA		1901
QY	129	ACAAGGGAGTCTCTAGCAGCGCTGCCTGACATCATTTAAACC-----GTGCTGCG		178
Db	1902	ACACTAGCTCTTAGCGTGGCGCGCTGCCAATCATTTAACTCTGTGCAAGTGGCGGG		1966
QY	179	AGCGAGCCCTTCATTAAAGGCCCTGGGTATGGCCAGCCAGCATGTGTCACTGCCCGCGAGA		238
Db	1962	CGTGTGCCCTTTTAAAGTGCGGCTGTGTTCAGGAGCATGGCCACCGCGATCC----		2011
QY	239	CACAAACCCAGCAGCATTTGAACACTGC-ACAAGGCCATCTGGCCAGAGACTGTGACCA		297
Db	2018	-----CATTCAGGCAAGCATCTGGCGCGCGCGCGCCGACACCTCCCAAGAGACATCTGGCCA		2071
QY	298	CCACTTTC	304	
Db	2074	CGGCTCC	2080	

```

RESULT 14
US-10-893-315-160
; Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397

```

/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 2172
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 160
/ LENGTH: 9174
/ TYPE: DNA
/ ORGANISM: Human
US-10-893-315-160

Query Match 19.9%; Score 63.8; DB 22; Length 9174;
Best Local Similarity 63.2%; Pred. No. 4.3e-11;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 69 TCCAGGAACAGGAGTSCCCCGTGGCCCATGTACAGCAAAACAACTATTTTGA 128
DB 1842 TCCTCCACAGCGGTTTCGTAGGGCCCTTGGGTGACAGCAAAACAACTATTTTGA 1901
QY 129 ACACGGGATCTAGACGCTGCCCTGACAAATCAATTAACCC-----GTGCTGCCG 178
DB 1902 ACACTCAGCTCTAGCGTGGCGCTGCCAATCAATTAACCTCTGTGCAAGTGGCGCG 1961
QY 179 AGCCAGCCCTTATAGGCGCTGGGTATGGCCAGCCAGATGTCCACTGCCCGCGAGA 238
DB 1962 CCTGTCCCTTTATAGGCGCGCTGTGTCCAGGAGCATGCGCCATCC----- 2017
QY 239 CACAACCCAGGAGCATTTGAACACTGC-ACAGGCCATCTGCCAGAGAGCTGTACCA 297
DB 2018 ----CATCCAGGAGCATCTGCCCGCGCGCCGCCACCTCCAGAGAGCATGGCCA 2073
QY 298 CCACTTC 304
DB 2074 CCGCTCC 2080

RESULT 15
US-10-388-934-36/c
/ Sequence 36, Application US/10388934
/ Publication No. US20040005547A1
/ GENERAL INFORMATION:
/ APPLICANT: Boess, Franziska
/ APPLICANT: Suter-Dick, Laura
/ TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
/ FILE REFERENCE: 21199
/ CURRENT APPLICATION NUMBER: US/10/388,934
/ PRIOR APPLICATION NUMBER: 02005336.9
/ PRIOR FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 02015657.6
/ PRIOR FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 862
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 36
/ LENGTH: 13011
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-36

Query Match 17.2%; Score 55.2; DB 17; Length 13011;
Best Local Similarity 88.2%; Pred. No. 4.2e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCCGTCGCGCCATGTACACTGGGGGCCAGATCCAGGAACCAAGGAGTCCCGGTG 94
DB 3088 TGCCACGGAAGCCTTTGACACTGGGGGCCAGATCCAGGAACCAAGGAGTCCCGGTG 3029
QY 95 CGCCCATG 102
DB 3028 CGCCCATG 3021

Search completed: September 1, 2005, 16:44:28
Job time : 944.893 secs

; Patent No. 6825336

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-36

Query Match          19.9%; Score 63.8; DB 4; Length 9174;
Best Local Similarity 63.2%; Pred. No. 1.3e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 69 TCCAGAACCCAGCGAGTGCCTCCCTGCGCCCATGTCACAGAGAAAACAATTATTGGA 128
DB 1842 TCCCTCCACACAGAGGTTGCGTAGGCGCTTGAGTGACAGAAAACAATTATTGGA 1901
QY 129 ACACGGGAGATCTAGACGCGTGCCTGACATCATTAACCC-----GTGCTGCGCG 178
DB 1902 ACACATCAGCTCTAGGCTCGGCGCGCTGCATATCATTAACCTCTGTCGAAATGCGCGG 1961
QY 179 AGCCAGCCCTTATTAAGCCCTGGGTGANGGCCAGCCAGCATGATCCATGCCCGCGAGA 238
DB 1962 CCGTGCCCTTATTAAGGTGCGCGCTGTGTCCAGCGAGATGCGCACCGCCATCC---- 2017
QY 239 CACAACCCAGGAGATGTAACATCTGC-ACAAGGCACTGCTCCAGAGAGCTGTGACCA 297
DB 2018 ----CATTCAGGAGATCTGCGCGCGCGCCGCCACCTTCCAGAGAGACTGGCCA 2073
QY 298 CCACTTC 304
DB 2074 CCGCTCC 2080

RESULT 3
US-08-945-140-1/c
; Sequence 1, Application US/08945140
; Patent No. 6309878
; GENERAL INFORMATION:
; APPLICANT: CHEN, Ruihuan
; APPLICANT: DOIRON, Bruno
; APPLICANT: KAHN, Axel
; TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
; TITLE OF INVENTION: VECTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mallesop 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,140
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/04558
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: WO PCT/FR96/00560
; FILING DATE: 12-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky Esq., Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: EX95002-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match          17.2%; Score 55.2; DB 3; Length 194;
Best Local Similarity 88.2%; Pred. No. 1.6e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCGTGCGCCCATGTGACCTGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGTG 94
DB 78 TGCCAACGGAAGCCTTGACACTGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGTG 19
QY 95 CGCCCATG 102
DB 18 CGCCCATG 11

RESULT 4
US-08-791-849A-14/c
; Sequence 14, Application US/08791849A
; Patent No. 591449
; GENERAL INFORMATION:
; APPLICANT: Makoto MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,849A
; FILING DATE: January 30, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
FEATURE:
NAME/KEY: intron
LOCATION: join(3219..3765, 3949..5916, 6009..6151,
LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
LOCATION: 9480..10162)
US-08-791-849A-14
```

```
Query Match 17.2%; Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.1e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 35 TGCCCCGTCGCGCCATGTACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 94
DB 3088 TCCCAACGAAAGCCTTGTATGCTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 3029
QY 95 CGCCCATG 102
DB 3028 CGCCCATG 3021
```

```
RESULT 5
US-09-949-016-11786/c
Sequence 11786, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11786
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-11786
```

```
Query Match 12.5%; Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.011;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 34 GTGCCCCGTGCGCCATGTACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 93
DB 2417 GTTCCATTGGAAGCCTGTATGCTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357
```

```
RESULT 6
US-09-949-016-17205/c
```

```
Sequence 17205, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17205
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-17205
```

```
Query Match 12.5%; Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.011;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 34 GTGCCCCGTGCGCCATGTACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 93
DB 2417 GTTCCATTGGAAGCCTGTATGCTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357
```

```
RESULT 7
US-09-484-970B-138/c
Sequence 138, Application US/09484970B
Patent No. 6426186
GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmueth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 138
LENGTH: 3054
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Inocyte ID No. 6426186 336987.1CB1
US-09-484-970B-138
```

```
Query Match 11.4%; Score 36.6; DB 3; Length 3054;
Best Local Similarity 62.6%; Pred. No. 0.087;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 2 ACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGTGCCTCATGTACACTGGGG 61
DB 355 ACTTCGAAGCAAGACCCCTGACGACGGGTAGCTCGAGCGTCCATCCAAAGCCGGTG 296
QY 62 GCCAAGTCCAGAAACACGGAGTGCCTCGG 92
DB 295 GCTGAGGCTTGGAGCGAAGGAGAGCCGG 265
```

```
RESULT 8
```

US-09-949-016-19194/c
; Sequence 19194, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19194
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19194

Query Match 10.9%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGCTCCAGCCCGCCAGACA 240
DB 478 CCAGCCCTTCCTGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGAGTTAGGCC 419
QY 241 CAACCCGAGCAGCATTTGAACTGCACACGCGCATCTGCGAGAGCTGTGACCA 300
DB 418 CCCATCCAGAGACACATCCACCCGATCCACAGCCCTGCGACCAACCATTCATGATGCC 359
QY 301 CTT 303
DB 358 CTT 356

RESULT 9
US-09-949-016-73784/c
; Sequence 73784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73784
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-73784

Query Match 10.9%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGCTCCAGCCCGCCAGACA 240
DB 478 CCAGCCCTTCCTGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGAGTTAGGCC 419

QY 241 CAACCCGAGCAGCATTTGAACTGCACACGCGCATCTGCGAGAGCTGTGACCA 300
DB 418 CCCATCCAGAGACACATCCACCCGATCCACAGCCCTGCGACCAACCATTCATGATGCC 359
QY 301 CTT 303
DB 358 CTT 356

RESULT 10
US-09-949-016-2153/c
; Sequence 2153, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2153
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2153

Query Match 10.9%; Score 35; DB 4; Length 2393;
Best Local Similarity 55.3%; Pred. No. 0.26;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGCTCCAGCCCGCCAGACA 240
DB 2182 CCAGCCCTTCCTGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGAGTTAGGCC 2123
QY 241 CAACCCGAGCAGCATTTGAACTGCACACGCGCATCTGCGAGAGCTGTGACCA 300
DB 2122 CCCATCCAGAGACACATCCACCCGATCCACAGCCCTGCGACCAACCATTCATGATGCC 2063
QY 301 CTT 303
DB 2062 CTT 2060

RESULT 11
US-09-949-016-74/c
; Sequence 74, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 2408

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-74

Query Match
Best Local Similarity 10.9%; Score 35; DB 4; Length 2408;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 2190 CCAAGCCCTTCCTGGGGCGCTGTCTCTCTCCATCCAGCTTAGCCAGGTTAGGCC 2131

QY 241 CAAACCCAGCAGCATTTGAACATGCAAGCGGCATCTGCCAGAGAGCTGTGACCA 300
DB 2130 CCCATCCAGAGACATCCACCCGGATCCAGGCTTCCACCAACCATCTCATGATAGCC 2071

QY 301 CTT 303
DB 2070 CTT 2068

RESULT 12
US-09-949-016-11816/c
; Sequence 11816, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11816
; LENGTH: 7561
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11816

Query Match
Best Local Similarity 10.9%; Score 35; DB 4; Length 7561;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 5350 CCAAGCCCTTCCTGGGGCGCTGTCTCTCTCCATCCAGCTTAGCCAGGTTAGGCC 5291

QY 241 CAAACCCAGCAGCATTTGAACATGCAAGCGGCATCTGCCAGAGAGCTGTGACCA 300
DB 5290 CCCATCCAGAGACATCCACCCGGATCCAGGCTTCCACCAACCATCTCATGATAGCC 5231

QY 301 CTT 303
DB 5230 CTT 5228

RESULT 13
US-09-949-016-13895/c
; Sequence 13895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13895
; LENGTH: 7562
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13895

Query Match
Best Local Similarity 10.9%; Score 35; DB 4; Length 7562;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 5350 CCAAGCCCTTCCTGGGGCGCTGTCTCTCTCCATCCAGCTTAGCCAGGTTAGGCC 5291

QY 241 CAAACCCAGCAGCATTTGAACATGCAAGCGGCATCTGCCAGAGAGCTGTGACCA 300
DB 5290 CCCATCCAGAGACATCCACCCGGATCCAGGCTTCCACCAACCATCTCATGATAGCC 5231

QY 301 CTT 303
DB 5230 CTT 5228

RESULT 14
US-09-270-767-29006
; Sequence 29006, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 29006
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29006

Query Match
Best Local Similarity 10.4%; Score 33.4; DB 4; Length 364;
Matches 105; Conservative 0; Mismatches 96; Indels 4; Gaps 1;

QY 82 GAGATGCCCCCGTGGCCCATGTCAACAGCAAAACAACTTATTGTAACACAGGGGA---- 137
DB 160 GAGATGCTTATGCGGCCCATGTCAACAGTACCGCCATGATGTTCGACAAAGGGGACGGC 219

QY 138 TCCTAGCAGCGCTGCGCTGACATATTAACCGTGTCTGCGAGCGAGCCCTTCATTAAGGC 197
DB 220 GCCGAACAGACTTGTGTGCGATGCTACCGCCCTGTGATGATGTCCGCTATTATACGGGC 279

QY 198 CTTGGGTATGCGCAGCAGCATGATCACTGCCCCCGAGACACAAACCCAGCAGCATT 257
DB 280 CTTAACAGGGAAGAAAGAGCCAGACCACTAGCGGAAAAAACAACCCACCAACAT 339

QY 258 GAACATGACACAGCGCCATCTGCC 282
DB 340 GCAAAACAGCATTCAGCCACAGTCC 364

RESULT 15
US-09-270-767-13109
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 2316.24 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 taccactg999gscagagcc.....ttccgctactagctagccgc 372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214.4	57.6	1181	10	RAT1GFZ
2	214.4	57.6	185148	2	AC136382
3	188.8	50.8	1363	10	MMILGF
4	188.8	50.8	192843	10	AL607124
5	98.2	26.4	1569	10	BC013345
6	79.2	21.3	1507	10	AY560836
7	76.4	20.5	1500	6	AX401932
8	76.4	20.5	1500	6	AX827271
9	76.4	20.5	1500	10	RAT1GFB
10	76.4	20.5	5001	6	AX163782
11	76.4	20.5	5001	10	RAT1GFB
12	65.4	17.6	1510	10	BC078889
13	64	17.2	3886	9	AY095345
14	63.8	17.2	6128	6	AX409747
15	63.8	17.2	6128	9	HUM1GFBP1A
16	63.8	17.2	6480	11	G19994
17	63.8	17.2	6480	9	HUM1GFBP1
18	63.8	17.2	9082	9	AY434089
19	63.8	17.2	69887	9	AC091524

C 20	62.2	16.7	141539	9	AC146152	AC146152 Pan trogl
C 21	62.2	16.7	189932	2	AC146117	AC146117 Pan trogl
C 22	62.2	16.7	200935	2	AC148834	AC148834 Pan trogl
C 23	60.8	16.3	448	11	G67139	G67139 IGFBP1 X1.1
C 24	55.2	14.8	194	6	A57715	A57715 Sequence 1
C 25	55.2	14.8	194	6	AR175909	AR175909 Sequence
C 26	55.2	14.8	13011	6	E14395	E14395 gDNA encod
C 27	55.2	14.8	13011	6	AX827302	AX827302 Sequence
C 28	55.2	14.8	13011	10	RNLPG	RNLPG
C 29	54.4	14.6	895	9	HUM1GFBP1	X05684 Rat L-PK ge
C 30	52	14.0	539	11	G67173	M23592 Human insul
C 31	47.4	12.7	2717	10	RATPRL1	G67173 IGFBP1 5'A
C 32	47.4	12.7	231241	2	AC097039	M17088 Rat pyruvat
C 33	44.6	12.0	184689	9	AL442125	AC097039 Rattus no
C 34	44.6	12.0	256781	2	AC097952	AL442125 Human DNA
C 35	44.6	12.0	259329	2	AC109570	AC097952 Rattus no
C 36	44.2	11.9	1480	10	RAT1GFBP1	AC109570 Rattus no
C 37	44.2	11.9	206515	10	AC140332	M89791 Rattus norv
C 38	44.2	11.9	215533	2	AC149086	AC140332 Mus muscu
C 39	42.6	11.5	43295	2	AC006177	AC149086 Mus muscu
C 40	42.6	11.5	53370	9	AL592071	AC006177 Homo sapi
C 41	42.6	11.5	186314	10	AC003694	AL592071 Human DNA
C 42	42.4	11.4	170654	10	AC116502	AC003694 Mus muscu
C 43	42.4	11.4	187397	2	AC114566	AC116502 Mus muscu
C 44	41.4	11.1	200340	9	AC100797	AC114566 Mus muscu
C 45	41	11.0	81704	9	AC110299	AC100797 Homo sapi

ALIGNMENTS

RESULT 1
RAT1GFZ
LOCUS
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION M84484
VERSION M84484.1 GI:204927
KEYWORDS insulin-like growth factor.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae;

REFERENCE
1 (bases 1 to 1181)
Unterman,T.G., Iacson,R.G., McGary,E., Whalen,C. and Goswami,R.G.
AUTHORS Biochem. Biophys. Res. Commun. (1991) In press
JOURNAL Original source text: Rattus norvegicus (strain Sprague-Dawley)
COMMENT male adult liver DNA.

FEATURES
source
1..1181
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

Query Match 57.6% Score 214.4; DB 10; Length 1181;
Beet Local Similarity 97.3%; Pred. No. 1.5e-51;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 144 TGGCCCATGTCTACAGCAAACTATTGTGAACACGGGATCTTACAGCTGCC 203
DB 813 TGTGGCGAGCTCACAGAAAACAACTATTGTGAACACGGGATCTTACAGCTGCC 872
QY 204 CTGACATCATTAACCCGTCTGCCGAGCAGCCCTTCATAGGCGCTGAGTGGCCAG 263
DB 873 CTGACATCATTAACCCGTCTGCCGAGCAGCCCTTCATAGGCGCTGAGTGGCCAG 932
QY 264 CCAGCATGTCACATGCGCCGAGAGACCAACCCAGGACGATTTGAACACTGACACAGCG 323

Db 993 CCAGCATGGTCCAGTCCGCCGCCGAGACAAACCAGCAGCATGTGAACACTGCACACAGC 992
 QY 324 CCATCTGCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 367
 Db 993 CCATCTGCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 1036

RESULT 2
 AC136382
 LOCUS
 DEFINITION
 AC136382 185148 bp DNA linear HTG 01-NOV-2002
 Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
 *** 63 unordered pieces.
 AC136382
 AC136382.1 GI:24462257
 HTG: HTGS_PHASE1
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 185148)
 Muzny,D,Marie, Metzker,M,lee, Abranzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsdorfs,S, Amin,A, Anguiano,D, Anyalelechi,V, Ayagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Gueraa,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K, Harrey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howells,S, Huiyk,S, Hume,U, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshewa,L, Louiseged,H, Lozado,R,U, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhinney,S, McLeod,M, Mcneill,T, Meenen,E, Milosavljevic,A, Miner,G, Mijta,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Naiz,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwokeneme,O, Okonofu,G, Olariunsgoon,A, Pal,S, Parks,K, Paeternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, Puazo,M, Quiroz,J, Rachin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojka,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Saverly,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,U, Shvartbeyan,A, Sisson,I, Sitter,C,D, Smaaj,D, Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Suton,A, Svatek,A, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczek,R, Woodson,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhou,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Direct Submission
 Unpublished
 2 (bases 1 to 185148)
 Rat Genome Sequencing Consortium.

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 Center project name: KDGs
 Center Clone name: CH230-97018
 Summary Statistics
 Sequencing vector: Plasmid:
 Chemistry: Dye-terminator Big Dye 100k of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 135613 bases at least Q40
 Consensus quality: 140849 bases at least Q30
 Consensus quality: 145680 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently consists of 63 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1010: contig of 1010 bp in length
 1111: gap of unknown length
 1111: contig of 1498 bp in length
 2609: gap of unknown length
 2708: contig of 1207 bp in length
 3915: gap of unknown length
 4015: contig of 1532 bp in length
 4016: gap of unknown length
 5548: gap of unknown length
 5647: contig of 1472 bp in length
 7120: gap of unknown length
 7219: contig of 1666 bp in length
 7220: gap of unknown length
 8885: gap of unknown length
 8886: contig of 1273 bp in length
 8986: gap of unknown length
 10259: contig of 1273 bp in length
 10259: gap of unknown length
 10359: contig of 1268 bp in length
 11627: gap of unknown length
 11726: contig of 1762 bp in length
 13489: gap of unknown length
 13489: contig of 1534 bp in length
 13589: gap of unknown length
 15123: contig of 1724 bp in length
 15223: gap of unknown length
 16947: gap of unknown length
 17046: contig of 1118 bp in length
 18164: gap of unknown length
 18264: contig of 1414 bp in length
 19678: gap of unknown length
 19779: contig of 1909 bp in length
 21687: gap of unknown length
 21787: contig of 1105 bp in length
 21882: gap of unknown length
 22992: contig of 1344 bp in length
 24336: gap of unknown length
 24437: contig of 2181 bp in length
 2617: gap of unknown length
 26717: contig of 1860 bp in length
 28578: gap of unknown length
 28677: contig of 2146 bp in length
 30823: gap of unknown length
 30923: contig of 1343 bp in length
 30924: gap of unknown length
 32265: contig of 1528 bp in length
 32366: gap of unknown length
 32367: contig of 1528 bp in length
 33895: gap of unknown length
 33994: contig of 1379 bp in length
 35373: gap of unknown length

Db	99303	CTGACATCATCTTAAACCCTGGTCGTCCAGCGCAGCCCTTCAATAAGCCCTGGATTAGCCAG	99362
Oy	264	CCAGCATGTGTCACATGCCCCGCCGAGACACAACCCAGCAGCATTTGAACACTGCACACGG	323
Db	99363	CCAGCATGTGTCACATGCCCCGCCGAGACACAACCCAGCAGCATTTGAACACTGCACACGG	99422
Oy	324	CCATCTGCCAAGAGAGCTGTGACCAACCACTTCGCGTAATGACTA	367
Db	99423	CCATCTGCCAAGAGAGCTGTGACCAACCACTTCGCGTAATGACTA	99466
RESULT 3			
MWITGF			
LOCUS		1363 bp	DNA linear
DEFINITION		M.musculus gene for insulin-like growth factor binding protein-1.	ROD 01-AUG-1996
ACCESSION		X67493	
VERSION		X67493.1 GI:52699	
KEYWORDS		insulin-like growth factor binding protein-1.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Milne,V., Mohn,K.L., Bucan,M. and Taub,R.	
TITLE		Structure and localization of the IGFBP-1 gene and its expression during liver regeneration	
JOURNAL		Hepatology 19 (3), 656-665 (1994)	
MEDLINE		94164648	
PUBMED		7509771	
REFERENCE		2 (bases 1 to 1363)	
AUTHORS		Mohn,K.L., Waddle,J.R. and Taub,R.	
TITLE		Comparison of mouse and human IGFBP-1 genes reveals a potential insulin-responsive sequence and conservation of all intron/exon boundaries	
JOURNAL		Nucleic Acids Res.	
REFERENCE		3 (bases 1 to 1363)	
AUTHORS		Taub,R.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard Hughes Medical Institute, Clinical Research Bldg., Room 475, 422 Curie Boulevard, Philadelphia, PA 19104-6145, USA	
FEATURES		Location/Qualifiers	
source		1..1363	
		/organism="Mus musculus"	
		/mol_type="genomic DNA"	
		/strain="NIH Swiss"	
		/db_xref="taxon:10090"	
		/clone="GW2.1"	
		/cell_line="NIH 3T3"	
		/clntype="fibroblast"	
		/clone_lib="Genomic, lambda FIX II"	
		/dev_stage="embryo"	
		500..508	
		/note="AP-2 consensus site"	
		692..706	
		/note="insulin-responsive element"	
		720..741	
		/note="Caat box with APF,HNF and NF-E1 consensus sequences"	
		764..768	
		792..1336	
TATA_signal			
exon			

```
misc_feature      792
                  /note="mRNA cap site"
CDS               964 .. >1336
                  /codon_start=1
                  /product="insulin-like growth factor binding protein-1"
                  /protein_id="CA47832.1"
                  /db_xref="GI:52700"
                  /db_xref="GOA:P47876"
                  /db_xref="UniProt/Swiss-Prot:P47876"
```

```

/translacion="MPELTIVSWPFLILLSFQIGVAGAPQPHCAPCTAERLGLCE
PVPASCEISRPAQCGCCPTCALPMGAACGVATARCAQGISCRALPGEPRPLHALTRG
QGA5LPEPAAPATSTLFSQHE"

```

Query Match	50.8%	Score 188.8	DB 10	Length 1363
Best Local Similarity	90.2%	Pred. NO. 4.3e-44		
Matches 202; Conservative	0	Mismatches 22	Indels 0	Gaps 0

OY	144	TGCGCCCATGTGACAAAGCAAAACAACTATTATTGTAACA	CGGGGATTCCTAGACGCTGGC	203
Db	667	TGTGTAGAGCTCAACAAGCAAAACAACTTATTATTGAACA	CTGGGGCTCTTAGACAGCTGTGC	726
OY	204	CTGCAATCATTTAAACCCGTGCTGCAGGCAAGCCCTTCA	TAAAGCCCTGGGATGAGCCAG	263
Db	727	CTGACAAATCATTTAACTGTGCGCGCAAGCAGACCTCT	CATTAAAGCTCTGGGTACGACAG	786
OY	264	CCAGCATGCTCCACTGCCCCGCGAGACACAACAAACCG	ACGAGCATTTGAACATCTGCACACGG	323
Db	787	CCAGCATGTGTCACTGCCCCGCGAGAACACACACCG	ACGAGCATTTGAACACTGCACACGG	846
OY	324	CCATTGTGCCAGAGGCTGTGACCAACCACTTCCGTTACT	AGTGA 367	
Db	847	CCGCTGTGCCAGAGGCTGTGACCAACCACTTGCACACT	ACTATATGTA 890	

RESULT 4				
AL607124		192843 bp	DNA	linear
LOCUS				
DEFINITION	Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete sequence.			
ACCESSION	AL607124			
VERSION	AL607124.15	GI:20145926		
KEYWORDS	HTG.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1			
AUTHORS	Oliver, K.			
TITLE	Direct Submission			
JOURNAL	Submitted (11-Apr-2002)			
	Wellcome Trust Sanger Institute, Hinxton,			

COMMENT

one plasmid subclone or more than one M13 subclone, and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Bm, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: DBACe3.6.

FEATURES
SOURCE

```
/organism="Mus musculus"  
/mol_type="genomic DNA"  
/db_xref="taxon:10090"  
/chromosome="11"
```

```

/clone="RP23-20C9"
/clone_lib="RPCI-23"

```

Query Match	50.8%	Score 188.8;	DB 10;	Length 192843;
Best Local Similarity	90.2%;	Pred. No. 5.5e-44;		
Matches 202; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;

QY	144	GGGCGCCATGTCACAGAGAAAACAACCTTATTTGAAACAGGGGATCTCTAGACGCTGCC	203
Db	2419	TGTGTAGAGCTCACAGCAAAACAACCTTATTTTGAACAGGGGCTCTTAGACCGCTGCC	2478
QY	204	CTGACAATCATTTAAACCCGCTGCTGCAGACCAAGCCCTTCATTAAGGCGCTGGATATGGCAG	263
Db	2479	CTGACAATCATTTAACTGTGTGCGGACAGCCGCTTCATTAAGGCTCTGGGATAGCACAG	2538
QY	264	CCAGCATGTCACACTGCGCCGCGGAGACAAACCAGCAGACATTTGAACACTGCACACGG	323
Db	2539	CCAGCATGTCACACTGCGCCGCGGAGACACACCAGCAGACATTTGAACACTGCACACGG	2598
QY	324	CCATCTGCCGACAGAGCTGTACACACCACTTCGGCTAATTAAGCTA	367
Db	2599	CCGCTCTGCCGACAGAGCTGTGTACACCACTTCGCACTAATTAAGCTA	2642

RESULT 5
BC013345

LOCUS	1669 bp	mRNA	linear	ROD 30-JUN-2004
DEFINITION	Mus musculus inauulin-like growth factor binding protein 1, mRNA			
ACCESSION	(Cpna clone MGC:14075 IMAGE:4161889), complete cds.			
VERSION	BC013345			
KEYWORDS	BC013345.1	GI:15426482		
SOURCE	MGC.			
	Mus musculus (house mouse)			

REFERENCE
AUTHORS
Straubenberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,
1 (bases 1 to 1569)
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae; Mus.

TITLE	JOURNAL	PUBMED	REFERENCE
Altschul, S.F., Zeeberg, B., Bhetton, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, I., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Caminici, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McErmann, K.C., Malek, J.A., Gunnarsson, P.H., Richards, S., Wexler, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Girmwood, J., Schmutz, J., Myers, R., Buterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schercher, A., Schein, J.E., Jones, S.J. and Marz, M.A.	human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	12477932
2 (bases 1 to 1569)			

AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
DNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology
<http://www.systembiology.org>
 contact: amadan@systembiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MQC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMN at: <http://image.llnl.gov>
 Series: IRK Plate: 18 Row: 1 Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1.1569
 /organism="Mus musculus"
 /mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MQC:14075 IMAGE:4161889"

/tissue_type="liver, normal, 5 month old male mouse."

/clone_id="NCI_CGAP_L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene

1.1569
 /gene="Igfbp1"

/note="synonym: IGFBP-1"

/db_xref="LOCUSID:16006"

/db_xref="MGI:96436"

CDS

204..1022

/gene="Igfbp1"

/codon_start=1

/product="Insulin-like growth factor binding protein 1"

/protein_id="AAH13345.1"

/db_xref="GI:15426483"

/db_xref="MGI:96436"

/translation="MPEPLTVSWPRLILSFQIVAGAPQPMHCAPCAETARLGLCP
 PVPASCPESRPAAGCCCTCTCLPMGACGVAATARCAQGLSRALPGEPRPLHALTRG
 QGACVPEPATPTASGLSIEKEAKAVPVRVPESEKMTBEQLDSFHLMAFSED
 OPILMNAISTYSMPAREIADKKWKEPCQRELIVLERLAAQKAGEEIKFYLPLN
 CNKNGFYHSKQCEISLDEAEELCMCVPMWSGRIPGSLERDPRCHQFYVNVN"

ORIGIN

Query Match 26.4%; Score 98.2; DB 10; Length 1569;
 Best Local Similarity 92.8%; Pred. No. 1.2e-17;
 Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 257 TGGCCAGCAGCATGCTCCAGTCCCGCCGAGACAAACCGAGAGCATTAACACG 316
 |||
 DB 20 TGGGACGACGATGCTCCAGTCCCGCCGAGACAAACCGAGAGCATTAACACG 79
 |||
 QY 317 CACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 367
 |||
 DB 80 CACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 130
 |||

RESULT 6
 LOCUS AY560836 1507 bp mRNA linear ROD 22-MAR-2004
 DEFINITION Spermophilus tridecemlineatus insulin-like growth factor binding
 protein 1 (IGFBP1) mRNA, complete cde.
 ACCESSION AY560836
 VERSION AY560836.1 GI:45505308

SOURCE
 ORGANISM Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
 Spermophilus tridecemlineatus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scuriinae;
 Spermophilus.

REFERENCE
 1 (bases 1 to 1507)
 Li, Y., Klimants, D. and Hallenbeck, J.M.
 Cloning and characterization of insulin-like growth factor binding
 protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1507)
 AUTHORS Li, Y., Klimants, D. and Hallenbeck, J.M.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
 Bethesda, MD 20892, USA

FEATURES
 source location/Qualifiers

source

1.1507
 /organism="Spermophilus tridecemlineatus"

/mol_type="mRNA"

/db_xref="taxon:43179"

gene

1.1507
 /gene="IGFBP1"

689..1507
 /gene="IGFBP1"

/codon_start=1

/product="Insulin-like growth factor binding protein 1"

/protein_id="AA567029.1"

/db_xref="GI:45505309"

/translation="MPEVPAAGLMPRLILAVQVSTVASTQPMHCAPCSAKALCP
 PVPSCPELSRPAAGCCPCALPLGACGVAATATYAGLSLALPGEPRPLHALTRG
 QGACVPEPATPTASGLSIEKEAKAVPVRVPESEKMTBEQLDSFHLMAFSED
 OPILMNAISTYSMPAREIADKKWKEPCQRELIVLERLAAQKAGEEIKFYLPLN
 CNKNGFYHSKQCEISLDEAEELCMCVPMWSGRIPGSLERDPRCHQFYVNVN"

ORIGIN

Query Match 21.3%; Score 79.2; DB 10; Length 1507;
 Best Local Similarity 67.1%; Pred. No. 4e-12;
 Matches 149; Conservative 0; Mismatches 58; Indels 15; Gaps 2;

QY 144 TGGCCCATGTGCACAGAAACAACTATTGTAACAGCGGATCCTAGACGCTGCC 203
 |||
 DB 420 TGGCGAAGAACGAGAGCAAACTATTGTAACAGCGGATCCTAGACGCGCG 479
 |||
 QY 204 CTGACCAATCATTA-----CCCGTGTGCGCAGCCAGCCCTTCATAAGCGCTGG 253
 |||
 DB 480 CTGGCAATCATTACTTCTAGTCAAGTGAAGTAACCGGCCCTTTATAGCAAGG 539
 |||
 QY 254 GTATGCCAGCCAGCATGTCCTGACCTGCCCGCCGAGACAAACCGAGAGCATTAACA 313
 |||
 DB 540 CTGACCCAGCAGAGCATGTCGACCTGCCCGGAGCTCAGAAATTGACACTATCA-- 597
 |||
 QY 314 CTGACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCT 355
 |||
 DB 598 ---GCCACTGTCCAGCTGCCAGAGAGCATTAACCACTGTCC 636
 |||

RESULT 7
 LOCUS AX401932 1500 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 1608 from Patent WO0210453.
 ACCESSION AX401932
 VERSION AX401932.1 GI:21338112

SOURCE
 ORGANISM Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 Mendrick, D., Porter, M.W., Johnson, K.R., Casale, A.L. and
 Elashoff, M.R.
 TITLE Molecular toxicology modeling
 JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;
 Gene Logic, Inc. (US)

FEATURES
 source location/Qualifiers

source

1.1500
 /organism="Rattus norvegicus"

/mol_type="unassigned DNA"

/db_xref="taxon:10116"

/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

```

Query Match          20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 349
      |||
DB 8 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 8
AX827271 1500 bp DNA linear PAT 12-DEC-2003
LOCUS Sequence 5 from Patent EP1344834.
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
          F. HOFMANN-LA ROCHE AG (CH)
FEATURES
source 1..1500
        /organism="Rattus norvegicus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:10116"

ORIGIN
Query Match          20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 349
      |||
DB 8 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 9
RATIGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (RIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Mohm, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL is rapidly and highly induced in regenerating liver
MEDLINE Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
PUBMED 91141487
COMMENT 1705004
FEATURES
source Original source text: Rat, cDNA to mRNA.
        location/Qualifiers
        1..1500
        /organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AAA41380.1"
/db_xref="GI:204733"
/translation="MPEFLTVVSPFFILLSFQRYVYAGAPQPHCAPCTARELCP
PVPAACEIPISRPAGCGCCPTCALPLGACGATNRCQGLSCRLPEPPLHLATG
QGACVLEPPAAPATSLSGSOHEKXAAVASDELAESPEMTBEQLDSFILMAPSRD
QPIIMNAISTVSSMRAREITDLKKKEPCORELKYVERLAQAQKAGDEIYKFLPN
CNKRGFYHSKQCETSLDGBAGLCWCYVPMWGKIPGSLFETRGDYNCHQYFNQVQ"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match          20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 349
      |||
DB 8 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 10
AX163782 5001 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 46 from Patent WO0138579.
ACCESSION AX163782
VERSION AX163782.1 GI:14544878
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Gould-Rothberg, B.E., Dipippo, V.A., Rameeh, T.M. and Gerwein, R.W.
TITLE Method of identifying toxic agents using msaid-induced differential
JOURNAL gene expression in liver
PUBMED Patent: WO 0138579-A 46 31-MAY-2001;
        Curagen Corporation (US)
FEATURES
source 1..5001
        /organism="Rattus norvegicus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:10116"

ORIGIN
Query Match          20.5%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 2.8e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 349
      |||
DB 1 CACAAACCCGAGCATTGAACACTGCACACGGCCATTGCCCAAGAGCTGTGACCAC 60

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 61 CACTTCGGCTACTATCTA 78

RESULT 11
RATIGFB

```

LOCUS	LOCUS DEFINITION	5001 bp	DNA	linear	ROD 30-NOV-1995
ACCESSION	RATIGFBA				
VERSION	Rat insulin-like growth factor binding protein-1 complete cds.				
KEYWORDS	L22979.1 GI:1098472				
SOURCE	Insulin-like growth factor binding protein-1, <i>Rattus norvegicus</i> (Norway rat)				
ORGANISM	<i>Rattus norvegicus</i>				
REFERENCE	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; <i>Rattus</i> .				
AUTHORS	1 (bases 1 to 5001)				
TITLE	Lacson, R., Ohler, D., Yang, E., Goswami, R. and Unterreiner, T.				
JOURNAL	Dideoxy sequencing and structural analysis of the rat insulin-like growth factor binding protein-1 gene				
MEDLINE	Biochim. Biophys. Acta 1218 (1), 95-98 (1994)				
COMMENT	94250701				
	7514892				
	On Nov 30, 1995 this sequence version replaced gi:385167.				
	Original source text: <i>Rattus norvegicus</i> (strain Sprague-Dawley) DNA.				

FEATURES	
source	Location/Qualifiers
	1..5001
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/strain="Sprague-Dawley"
	/db_xref="taxon:10116"
	/sex="male"
	/tissue_type="liver"
	/dev_stage="adult"
	/tissue_lib="Clontech EMBL-3 SP6/7"
exon	60..525
	/number=1
5' UTR	60..152
gene	153..4221
	/gene="IGFBP-1"
	join(153..525,1850..2034,2874..3002,4090..4221)
	/gene="IGFBP-1"
CDS	/codon_start=1
	/product="insulin-like growth factor binding protein"
	/protein_id="AA82581.1"
	/db_xref="GI:1098473"
	/translation="MPEFLTVSWPFLILSFQVRVVGAPQPMHCAPTAERLELCP PVPSPCEISRPAGCGCCPTCALPLGAAGVATACAGLSCRALPEEPRIHATLGG QGCVLEPAPATSSLSGSGHEEKAAVASDELAESPMTREGLDLSFHMARESD OPTIMAINISTYSWRAREITDLKKWKEPCQELTKYLERLAAQOKADELYKFLPN CNKNGFPHSKCSEFLSDGEAGLCWCVYPWSGKKLPGSLETRGGPNCHQYFNVO"
intron	526..1849
	/gene="IGFBP-1"
	/number=1
exon	1850..2034
	/gene="IGFBP-1"
	/number=2
intron	2035..2873
	/gene="IGFBP-1"
	/number=2
exon	2874..3002
	/gene="IGFBP-1"
	/number=3
intron	3003..4089
	/gene="IGFBP-1"
	/number=3
exon	4090..4743
	/number=4
3' UTR	4222..4743
ORIGIN	
Query Match	20.5%; Score 76.4; DB 10; Length 5001;
Best Local Similarity	98.7%; Pred. No. 2,86-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
290	CACAAACCGACGATTAACACTGCACACGGCCATCTCCACAGAGAGTGTGACAC 349

Db	Accession	Definition	Version	Source	Organism
Db	1	CACAAACCCAGCAGCATTTGAACACTCCACACGCGCATCTGCCAGAGAGCTGTGACCAC	60		
Qy	350	CACCTCCGCTACTACTGTA 367			
Db	61	CACCTCCGCTACTACTGTA 78			
RESULT 12					
LOCUS	BC078889	1510 bp mRNA linear ROD 03-AUG-2004			
DEFINITION		Rattus norvegicus insulin-like growth factor binding protein 1, mRNA (CDNA clone MGC:93595 IMAGE:7129185), complete cds.			
ACCESSION	BC078889				
VERSION	BC078889.1	GI:50927646			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
JOURNAL					
PUBLISHED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
REMARK					
COMMENT					

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 CDNA Library Preparation: Express Genomics
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-ehgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 184 Row: e Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6981079.
 Location/Qualifiers
 1. 1510
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"

gene

CDS

ORIGIN

Query Match
Best Local Similarity 98.5%; Pred. No. 4.3e-08;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 301 CGAGCATTTGAACACTGTCACACGGCCATCTGCCAGAGAGCTGTGACCACCACTTCCGCTA 360
DB 1 CGAGCATTTGAACACTGTCACACGGCCATCTGCCAGAGAGCTGTGACCACCACTTCCGCTA 60

OY 361 CTAGCTA 367
DB 61 CTATCTA 67

RESULT 13
LOCUS AY095345 3886 bp DNA linear PRI 30-DEC-2002
DEFINITION Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
ACCESSION AY095345
VERSION AY095345
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 3886)
Kim,J.-J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
Regulation of insulin-like growth factor binding protein-1 promoter
activity by PCHR and HOXA10 in primate endometrial cells
Biol. Reprod. 68 (1), 24-30 (2003)
12493691
12493691
2 (bases 1 to 3886)
Kim,J.-J., Jaffe,R.C. and Fazleabas,A.T.
Direct Submission
Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

JOURNAL
REFERENCE
AUTHORS
TITLE
PUBMED
REFERENCES
JOURNAL

FEATURES
source

gene

Location/Qualifiers
1..3886
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/gene="IGFBP-1"

	mRNA	<3655..>3886 /gene="IGFBP-1" /product="insulin-like growth factor binding protein-1" 3655..>3886 /gene="IGFBP-1" /codon_start=1 /product="insulin-like growth factor binding protein-1" /protein_id="AA02373.1" /db_xref="GI:20853765" /translation="MSEVPARVWLVELLITVGQVTSAPWOCAPCSAEKIALCPY" PASCEVTBSAGCCGCPMCALPLGAACGVATAR"
CDS		
ORIGIN		
Query Match	17.2%; Score 64;	DB 9; Length 3886;
Best Local Similarity	60.6%; Pred. No. 1.1e-07;	
Matches 149; Conservative	0; Mismatches 80; Indels 17; Gaps 2.	
OY	120 TCCAGGAAACAAGGGAGTGCCTCGGCCCATGTCAACAAGCAAACACTTAATTGGA	179
Db	3331 TCCTCCCACTCGCGGTTTGCGTAGGGCCTTGGCGCACTACAAAACAACTTAATTGGA	3390
OY	180 ACACGGGGATCCTAGACAGCTGCCTCGTCACATCAATTAACC-----GTGCTGCCG	229
Db	3391 ACATCAGACTCTTAGCCGCGCGCTGCCAATCAATTAACCTCTGTGTCAGATGGCGCGG	3450
OY	230 AGCCAGCCCTTCATPAAGGCCCCGTGGGTATGACCAAGCAGACTGTCCACTGCCCCGCGAGA	289
Db	3451 CCTGAGCCCTTATPAAGGCGCGCGCTGTGTCAGAACAGACTGGGCAAGCCATCCATC	3510
OY	290 CACAAACCCAGCAGACTTGAACACTGACACACGCGCCATCTGCCAGAGAGCTGACAC	349
Db	3511 CAGCAAGC-----ATCTGCCGCGCGCGCGCCACCTTCCAGAGAGCACTGGCCAC	3563
OY	350 CACTTC 355	
Db	3564 CGCTCC 3569	
RESULT 14		
LOCUS	AX409747	6128 bp DNA linear PAT 14-JUN-2002
DEFINITION	Sequence 2394 from Patent WO0229103.	
ACCESSION	AX409747	
VERSION	AX409747.1 GI:21442452	
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1	
AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 2394 11-APR-2002;	
FEATURES	GENE LOGIC INC (US) Location/Qualifiers 1..6128 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="EMBL/GenBank Accession No. M74587"	
ORIGIN		
Query Match	17.2%; Score 63.8; DB 6; Length 6128;	
Best Local Similarity	63.2%; Pred. No. 1.3e-07;	
Matches 156; Conservative	0; Mismatches 72; Indels 19; Gaps 3;	
OY	120 TCCAGGAACAAGGGAGTGCCTCGGCCCATGTCAACAAGCAAACACTTAATTGGA	179
Db	438 TCCTCCCACTCGCGGTTTGCGTAGGGCCTTGGCGCACTACAAAACAACTTAATTGGA	497
OY	180 ACACGGGGATCCTAGACAGCTGCCTCGTCACATCAATTAACC-----GTGCTGCCG	229
Db	498 ACATCAGACTCTTAGCCGCGCGCTGCCAATCAATTAACCTCTGTGTCAGATGGCGCGG	557

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 2828.19 Seconds
(without alignment)
5006.706 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 tacactggggggccagagatcc.....ttcgcctactagctagccgc 372

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_esc1: *
2: gb_esc2: *
3: gb_hic: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	26.4	1013	2	BF236974 602026605
2	87.8	22.6	310	6	BY783538 BY783538
3	87.8	22.6	310	6	BY794229 BY794229
4	87.8	22.6	354	6	BY771317 BY771317
5	86.8	23.3	401	1	AI785818 u17805.y
6	86.8	23.3	480	1	AI196314 u171a07.y
7	86.8	23.3	706	1	AI530146 u189f09.y
8	86.8	23.3	765	1	AI098594 uel1e07.y
9	86.8	23.3	789	1	AI530313 u191f01.y
10	86.8	23.3	795	1	AI529939 u187c09.y
11	86.8	23.3	811	6	CA478518 AGENCOURT
12	86.8	23.3	848	1	AI790802 u128b10.y
13	86.8	23.3	852	1	AI528304 u195g10.y
14	85.2	22.9	605	1	AI196154 u169d08.y
15	82.2	22.1	380	1	AI785039 u173a06.y
16	81.4	21.9	846	7	COS73026 AGENCOURT
17	80.4	21.6	692	7	CV127049 AGENCOURT
18	79.4	21.3	713	7	COS60662 AGENCOURT
19	75.4	20.3	748	7	COS75629 AGENCOURT
20	73.8	19.8	694	7	CV117001 AGENCOURT
21	72.4	19.7	632	7	W30013 mc24c07.x1
22	72.4	19.5	618	6	BB660958 BB660958
23	72.4	19.5	618	6	CD561711 B0435D02-
24	72.4	19.5	632	1	AI892189 mj88b01.y

25	67.4	18.1	488	1	AA674302 VP96g10.r
26	65.4	17.6	801	7	CK472246 AGENCOURT
27	64.4	17.3	840	7	CK473709 AGENCOURT
28	62.8	16.9	759	1	AA105355 mp37d09.r
29	61.8	16.6	269	2	BB604790 BB604790
30	61.6	16.6	587	2	AW916227 EST347531
31	60.8	16.3	1589	3	CR621807 full-1eng
32	60.8	16.3	1601	3	CR595377 full-1eng
33	59	15.9	545	1	AA060360 mj67a12.r
34	47.8	12.8	1159	8	CNS015XR AL106041 Drosophila
35	47.2	12.7	659	9	AZ840793 2M0138D02
36	44.4	11.9	1128	3	CR657703 Tetradon
37	44.4	11.9	1128	3	CR671810 Tetradon
38	44.4	11.9	1159	3	CR663312 Tetradon
39	44.4	11.9	1167	3	CR667848 Tetradon
40	42.2	11.3	597	8	AZ652514 IM0525H14
41	41.6	11.2	325	2	AW855818 RCI-CT027
42	40	10.8	212	2	BR149547 RCI-HT025
43	39.4	10.6	500	8	AO612859 HS-5116.A
44	39.2	10.5	234	2	AW886850 RCI-OT008
45	39.2	10.5	289	2	BE066031 RCI-BT031

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:416189 5',
BF236974 mRNA sequence.

ACCESSION BF236974 GI:11150891
VERSION BF236974.1
KEYWORDS EST.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE 1 (bases 1 to 1013)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence atp: 581.

FEATURES

source

1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:416189"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-Sport6; site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 26.4% Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 5.9e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

CY 257 TGGCAGCGACGATGCTCCTGCGCCGCGAGACAAACGACGACATTGAACACTG 316

Db	9	TGGGACGCGAGATGATGTCACCTGCGCGGAGACACACACCCAGCGAGATTGAACCTG	68
Oy	317	CACACGCGCCATCTGCGCGAGAGAGCTGTGTGACACCACTTCGCTACTAGCTA	367
Db	69	CACACGCGCGTCTGCGCGAGAGAGCTGTGTGACCACTTATGCACTACTATCTA	119
RESULT 2			
LOCUS	BY783538		
DEFINITION	BY783538 RIKEN full-length enriched, 17.5 days embryo whole body		
ACCESSION	Mus musculus cDNA clone U930176D05 5', mRNA sequence.		
VERSION	BY783538		
KEYWORDS	BY783538.1 GI:39710177		
SOURCE	EST.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 310) Carninci,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M., Aizawa,K., Arikawa,T., Ishii,Y., Sasaki,D., Sato,H., Kondo,S., Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watabiki,A., Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A., Kikuchi,N., Yoshiki,A., Kusakabe,M., Guscincich,S., Beisel,K., Pavan,W., Aldrich,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T., Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagiolini,M., Hensch,T.K., Brinkmeier,M., Camper,S., Hiroca,Y., Mombaerts,P., Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y. Targeting a complex transposome: the construction of the mouse full-length cDNA encyclopedia Genome Res. 13 (6B), 1273-1289 (2003)		
TITLE	full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES			
source	Location/Qualifiers 1..310 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="U930176D05" /tissue_type="whole body" /dev_stage="17.5 days embryo" /clone_id="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	23.6%	Score 87.8; DB 6; Length 310;	
Best Local Similarity	92.9%	Pred. No. 4.2e-14;	
Matches	92; Conservative	0; Mismatches 7; Indels 0; Gaps 0	
Oy	269	ATGTGTCACCTGCGCGCGAGACACAAACCCAGCGAGATTGAACACTGACACGCGCATC	328
Db	2	ATGTGTCACCTGCGCGCGAGACACACACCCAGCGAGATTGAACACTGACACGCGCGTC	61
Oy	329	TGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA	367

DB 62 TGCCAGAGAGCTGTGACCAACCACTTGCCACATCATCTCA 100

RESULT 3
BY794229

LOCUS
BY794229

DEFINITION
BY794229 RIKEN full-length enriched, 17.5 days embryo whole body

ACCESSION
BY794229

VERSION
BY794229.1

KEYWORDS
GI:39720868

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 331)

AUTHORS
Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arawaka, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Suohara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanabe, A., Hirozane-Kitahara, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustinich, S., Belsel, K., Pavan, W., Aldini, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

TITLE
Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia

JOURNAL
Genome Res. 13 (6B), 1273-1289 (2003)

MEDLINE
22703353

PUBMED
12819125

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

FEATURES
source
location/Qualifiers
1..331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930292H24"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_id="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN
Query Match 23.6%; Score 87.8; DB 6; Length 331;
Best local Similarity 92.9%; Pred. No. 4.3e-14;
Matches 92; Conservativity 0; Mismatches 7; Indels 0; Gaps 0;

DB 329 TGCCAGAGAGCTGTGACCAACCACTTGCCACATCATCTCA 367

DB 62 TGCCAGAGAGCTGTGACCAACCACTTGCCACATCATCTCA 100

RESULT 4
BY771317

LOCUS BY77317 354 bp mRNA linear EST 23-MAR-2004
DEFINITION BY77317 RIKEN full-length enriched, 17.5 days embryo whole body
MUS musculus cDNA clone L930052G15 5', mRNA sequence.
ACCESSION BY77317
VERSION BY77317.1 GI:39697955
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Suganara, Y., Saito, R., Oosato, N., Fukuda, S., Sato, K., Matshiki, A., Hirozane-Kitahara, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustinich, S., Betzel, K., Pavan, W., Aldinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakachi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
TITLE Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia
JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)
MEDLINE 22703353
PUBMED 12819125
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/ CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.
FEATURES
source Location/Qualifiers
1..354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930052G15"
/issue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_1fb="RIKEN full-length enriched, 17.5 days embryo whole body"
ORIGIN
Query Match 23.6%; Score 87.8; DB 6; Length 354;
Best Local Similarity 92.9%; Pred. No. 4.3e-14;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 269 ATGTCTCACTGCCCGCGGAGACACAAACCCAGGACATTGAACACTGACACAGGCCATC 328
DB 2 ATGTCTCACTGCCCGCGGAGACACACACCCAGGACATTGAACACTGACACAGGCCGCTC 61
QY 329 TGCCGAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 367
DB 62 TGCCGAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 100
RESULT 5
LOCUS A1785818 401 bp mRNA linear EST 02-UTL-1999
DEFINITION u178h05.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1887828 5' similar to gb:XB1579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION A1785818

VERSION A1785818.1 GI:5333534
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The MASHU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: u178h05.x1
Contact: Marra M/Waashu-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu
This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:972893
Seq primer: custom primer used
High quality sequence scop: 126.
FEATURES
source Location/Qualifiers
1..401
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1887828"
/sex="female"
/dev_stage="adult"
/lab_host="DHL10B"
/clone_1fb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME185-F13; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG]; digested and cloned into distinct DraIII sites of the pME185-F13 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAGCTGG and 3' end primer CGACTGCACTCGACGACA."
ORIGIN
Query Match 23.3%; Score 86.8; DB 1; Length 401;
Best Local Similarity 92.9%; Pred. No. 8.5e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 270 TGTCTCACTGCCCGCGGAGACACAAACCCAGGACATTGAACACTGACACAGGCCATCT 329
DB 1 TGTCTCACTGCCCGCGGAGACACACACCCAGGACATTGAACACTGACACAGGCCGCTC 60
QY 330 GCCGAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 367
DB 61 GCCGAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 98
RESULT 6
LOCUS A1196314 480 bp mRNA linear EST 14-OCT-1998
DEFINITION u171a07.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1887828 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:XB1579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION	AI196314	GI:3748920
VERSION	AI196314.1	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 400)	
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T. Gessel,S., Kucaba,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
TITLE	The WashU-HHMI Mouse EST Project	
JOURNAL	Unpublished (1995)	
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LINDL ; contact the IMAGE Consortium (info@image.jnl.gov) for further information. MGI:972152 Seq primer: custom primer used High quality sequence scop: 375. Location/Qualifiers 1..480	
FEATURES		
Source		

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887828"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone.lib="Sugano mouse liver ml1a"
/note="Organ: liver; Vector: pME18s-FL3; Site_1: DraIII
(CACGATG) ; Site_2: DraIII (CACCATG) ; 1st strand cDNA
was primed with an oligo(dT) primer
(ATGGGCGCTTTTATTTTTTTTTT) ; double-stranded cDNA was
ligated to a DraIII adaptor (TTGGGCGCCACGCG) , digested
and cloned into distinct DraIII sites of the pME18s-FL3
vector (5' site CACGATG, 3' site CACCATG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTTAAAGCTGCG and 3' end
primer GGACCTGACGCTGAGCA."

```

		23.3%	Score 86.8	DB 1:	Length 480;
Query Match		92.9%	Pred No. 8.8e-14,		
Best Local Similarity		91; Conservative	0; Mismatches 7;	Indels 0;	Gaps 0
Matches					
QY	270	TGTTCCACTGCCCGCCGAGACACAACCACCGAGCACTTGACACGGCCATCT	329		
Dd	1	TGTGTCACTGCCCGCGGAGACACACCCACGGACATTGAACAATGCACACGGCCGCTT	60		
QY	330	GCCCCAGAGAGCTGTGACCAACCATTCCTCCGCTACTAGCTA	367		
Dd	61	GCCCAGAGAGCTGTGACCAACCATTCGCCACTACTATCTA	98		
RESULT 7					
AIS30146					
LOCUS					
DEFINITION	AIS30146	706 bp	mRNA	linear	EST 18-MAR-1999
FACTORS	U89EF09.Y1	Sugano mouse liver	mla nus	musculus cDNA clone	
IMAGE	1889609.5'	similar to gb:M59316 rat	INSULIN-LIKE GROWTH		
FACTOR BINDING	PROTEIN 1	PRECURSOR (HUMAN);	gb:X81579	M.musculus	
mRNA	for insulin-like growth factor binding	(MOUSE);	mRNA		

ACCESSION	sequence.
VERSION	AT530146
KEYWORDS	AT530146.1 GI:44444281
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 706)
TITLE	Marrs,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ratter,E., Korn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
JOURNAL	The WashU-NCI Mouse EST Project 1999
COMMENT	Unpublished (1999) Contact: Maira M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:973933 Seq primer: custom primer used High quality sequence scop: 479.
FEATURES	
Source	1..706 Location/Qualifiers

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889609"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver m1a"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACCTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
(ATGGGCGCTTTTCTTTTCTTTT); double-stranded cDNA was
ligated to a DraIII adaptor (TTGTTGGCGTCTGTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAAGCTGCG and 3'
primer GCACCTGCAGCTCGAGCA."

```

Query Match	23.3%	Score 86.8	DB 1	Length 706
Best Local Similarity	92.9%	Pred. No. 9.5e-14		
Matches 91	Conservative 0	Mismatches 7	Indels 0	Gaps 0
QY	270	TGTTCCACTGCGCCGCCGAGACACAAACCAGCGAGCATTTGAACCTGCACACGCCCATCT	329	
Db	1	TGTTCCACTGCGCCGCCGAGACACACACCCAGCGAGCATTTGAACCTGCACACGCCCGCTCT	60	
QY	330	GCCCGAGAGGCTGTGACCAACCACTTCCGCTACTAGCTA	367	
Db	61	GCCCGAGAGGCTGTGACCAACCACTTCCGCTACTACTATCTA	98	
RESULT 8				
LOCUS	A1098594	765 bp	linear	EST 20-AUG-1999
DEFINITION	u031607.y1 Sugano mouse liver m1a Mus musculus cDNA clone IMAGE:1481988 5' similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.			
ACCESSION	A1098594			

```

VERSION      AI098594.1  GI:3448119
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 765)
Marra,M., Hillel,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thaising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra W/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LBNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 335.
Location/Qualifiers
1..765

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1481988"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_idb="Sugano mouse liver m1a"
/notes="Organ: liver; Vector: pMHS-F13; Site_1: DraIII
(CACGTGTG) ; Site_2: DraIII (CACCATGT) ; 1st strand cDNA
was primed with an oligo (drr) primer
(ATGTGGCCCTTTTCTTTTTTTTTT) ; double-stranded cDNA was
ligated to a DraIII adaptor [GTTGGCCTACTGG] , digested
and cloned into distinct DraIII sites of the pMHS-F13
vector (5' site CACGTGTG, 3' site CACCATGT). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCCTCTTAAGACTGG and 3' end
primer CGACCTCAGCTCGAGACA."

```

ORIGIN	Query Match	23.3%	Score 86.8	DB 1	Length 765
Best Local Similarity	92.9%	Pred. No. 9.7e-14			
Matches	91	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0
OY	270	TGTTCCACTGCCCGCCGAGACAAACCAGGAGACTTGAACACTGACACAGGCACTCT	329		
Db	1	TGTTCCACTGCCCGGAGAGACACACCCAGCAGCACTTGAACAACCTGACACAGGCGCTCT	60		
OY	330	GCCCGAGAGCTGTGACACCACTTCCGCTACTAGCTA	367		
Db	61	GCCCGAGAGCTGTGACCACTTCCGCTACTACTTA	98		
RESULT 9					
LOCUS	A1530313				
DEFINITION	A1530313	785 bp	mRNA	linear	EST 18-MAR-1999
	u191f01.v1	Sugano mouse liver	m14	mus musculus	cDNA clone
	IMAGE:1869785.5	' similar to gb:M59316	trial	INSULIN-LIKE GROWTH	
	FACTOR BINDING PROTEIN 1	PRECURSOR (HUMAN);	gb:X81579	M. musculus	
	mRNA for insulin-like growth factor binding	(MOUSE);	mRNA		
	sequence.				

ACCESSION A1530313
 VERSION A1530313.1 GI:4444448
 KEYWORDS EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 785)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Person,B., Shaller,T., Gibbons,M., Page,D., Harvey,N., Shuck,R.,
 Raltee,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LINT ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence, stop: 459.
 Location/Qualifiers
 1..785

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mRNA"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TCGTGGCTCTCG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAAGCTCG and 3' end
primer CGACCTGCACCTGAGACA."

```

ORIGIN	Query Match	23.3%	Score 86.8	DB 1	Length 785
Best Local Similarity	92.9%		Pred. No. 9.7e-14		
Matches	91	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0
QY	270	TGCTCCACTGCCCGCCGAGACACAAACCAGCGAGATTGAACACTGCACACGGCCACTT	329		
Db	1	TGGTCCACTCTCCCGCGGAGACACACACCCGCGGAGCTTGAACACTGCACACGGCCGCTCT	60		
QY	330	GCCGAGAGACTGTGACACCACTCTCGGCTACTGACTA	367		
Db	61	GCCGAGAGACTGTGACCACTCTCGGCTACTGACTA	98		
RESULT 10					
AI529939					
LOCUS					
DEFINITION	AI529939	799 bp	mRNA	linear	EST 18-MAR-1999
	U187C09.y1	Sugano mouse liver	mla	musculus	CDNA clone
	IMAGE:1889332.5	B1mlar	to gb:M59316	trial	INSULIN-LIKE GROWTH
	FACTOR BINDING PROTEIN 1	PRECURSOR (HDMAN)	gb:X81579	M. musculus	
	mRNA	for insulin-like growth factor binding (MOUSE);			mRNA
	sequence.				

ACCESSION A1529939 GI:4444074
 VERSION A1529939.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 799)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WaehU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 FEATURES
 source location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:189392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCAGTGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGCTGTGTGTG], digested and cloned into distinct DraIII sites of the pME18-FL3 vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGTCTCTTAAAGCTGCG and 3' end primer GCACCTGCAGCTCGAGCACA."

ORIGIN
 Query Match 23.3%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 9.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 270 TGGTCACCTGCGCCGCGAGACACAACCCAGGAGCATTTGAACACTGCACCGGCATCT 329
 |||||
 Db 1 TGTTCACCTGCGCCGCGAGACACAACCCAGGAGCATTTGAACACTGCACCGGCCTCT 60
 |||||
 Oy 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
 |||||
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTA 98
 |||||
 RESULT 11
 CA478518 811 bp mRNA linear EST 09-MAR-2004
 LOCUS AGENCOURT_10789306 NIH_MGC_152 Mus musculus cDNA clone
 DEFINITION IMAGE:6766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LKMO0138 row: m column: 23
 High quality sequence stop: 536.
 FEATURES
 source location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCCTGAGCGCCGCGACACACCTTGAAGAAGCTGGGTCTTTTCTTTTCTTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen. Library amplification 16 cycles. Library constructed by Mark Bittinger in the Bradfield Laboratory (McArdle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 23.3%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 9.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 270 TGGTCACCTGCGCCGCGAGACACAACCCAGGAGCATTTGAACACTGCACCGGCATCT 329
 |||||
 Db 39 TGTTCACCTGCGCCGCGAGACACACCCAGGAGCATTTGAACACTGCACCGGCCTCT 98
 |||||
 Oy 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
 |||||
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTA 136
 |||||
 RESULT 12
 A1790802 848 bp mRNA linear EST 02-JUL-1999
 LOCUS uk28b10.Y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 DEFINITION IMAGE:1970229 5', similar to gb:X8159.M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 848)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28b10.x1
Contact: Maira M/Washu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsoln.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 514.
Location/Qualifiers

1..848

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mklia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCAGTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTTCTTTT]
ligated to a DraIII adaptor (TGTGGCTTACTGTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGG and 3' end
primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match

Best Local Similarity 23.3%; Score 86.8; DB 1; Length 848;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGTTCACATGCGCGCGAGACACAAACCGAGAGATGAACTGACACAGGCGCATCT 329

Db 1 TGTTCACATGCGCGCGAGACACACCCAGAGATGAACTGACACAGGCGCATCT 60

QY 330 GCCCAGAGAGCTGTGTACCACTTCGCTACTAGCTA 367

Db 61 GCCCAGAGAGCTGTGTACCACTTCGCTACTACTACTA 98

RESULT 13

LOCUS

A1528304 852 bp mRNA linear EST 18-MAR-1999

DEFINITION

u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1890210.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION

A1528304

VERSION

A1528304.1 GI:4442439

KEYWORDS

EST.

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

Maira,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., Mcann,R.,
Waterston,R. and Wilson,R.
The Washu-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Maira M/Washu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsoln.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 478.
Location/Qualifiers

1..852

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCAGTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTTCTTTT]
ligated to a DraIII adaptor (TGTGGCTTACTGTG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGG and 3' end
primer CGACCTGCAGCTCGAGCACA."

ORIGIN

Query Match

Best Local Similarity 23.3%; Score 86.8; DB 1; Length 852;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGTTCACATGCGCGCGAGACACAAACCGAGAGATGAACTGACACAGGCGCATCT 329

Db 1 TGTTCACATGCGCGCGAGACACACCCAGAGATGAACTGACACAGGCGCATCT 60

QY 330 GCCCAGAGAGCTGTGTACCACTTCGCTACTAGCTA 367

Db 61 GCCCAGAGAGCTGTGTACCACTTCGCTACTACTACTA 98

RESULT 14

LOCUS

A1196154 605 bp mRNA linear EST 14-OCT-1998

DEFINITION

u195d08.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1887663.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION

A1196154

VERSION

A1196154.1 GI:3748760

KEYWORDS

EST.

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

Maira,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The Washu-HHMI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:971987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

```
1.605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG]; digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTCTCTTAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 22.9%; Score 85.2; DB 1; Length 605;
Pred. No. 2.6e-13; Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 270 TGTGTCACCTGCCCGCGAGACACCAACCCAGCAGCATTTGAACACTGCACAGCCCATCT 329
|||||
1 TGTGTCACCTGCCCGCGAGACACACACCAGCAGCATTTGAACACTGCACAGCGTCTCT 60

Db 1 TGTGTCACCTGCCCGCGAGACACACACCAGCAGCATTTGAACACTGCACAGCGTCTCT 60

Qy 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
|||||
61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98

Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98

RESULT 15
AI785039 380 bp mRNA linear EST 02-JUL-1999
LOCUS u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone

DEFINITION IMAGE:1888018.5, similar to gb:X81579.M.musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.
AI785039

ACCESSION AI785039.1 GI:5332755
VERSION EST.
KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 380)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pepe,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCam,R.,
Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

```
1.380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG]; digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTCTCTTAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 22.1%; Score 82.2; DB 1; Length 380;
Pred. No. 1.7e-12; Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 273 TCCACTGCCCGCGAGACACCAACCCAGCAGCATTTGAACACTGCACAGCCCATCTGCC 332
|||||
1 TCCACTGCCCGCGAGAGACACACACCAGCAGCATTTGAACACTGCATACGCCCGTCTGCC 60

Db 1 TCCACTGCCCGCGAGAGACACACACCAGCAGCATTTGAACACTGCATACGCCCGTCTGCC 60

Qy 333 CAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
|||||
61 CAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 95

Db 61 CAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 95

Search completed: September 1, 2005, 06:56:52
Job time : 2831.35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 1093.43 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916b-5
Perfect score: 372
Sequence: 1 taccactggggccagatcc.....ttccgctactagtagcgcgc 372

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372	100.0	372	US-09-972-916A-5	Sequence 5, Appli
2	321	86.3	321	US-09-972-916A-4	Sequence 4, Appli
3	223.6	60.1	423	US-09-972-916A-6	Sequence 6, Appli
4	219.4	59.0	270	US-09-972-916A-3	Sequence 3, Appli
5	219	58.9	219	US-09-972-916A-2	Sequence 2, Appli
6	154	41.4	423	US-09-972-916A-6	Sequence 6, Appli
7	76.4	20.5	1500	US-09-917-800A-1608	Sequence 1608, Ap

8	76.4	20.5	1500	17	US-10-388-934-5	Sequence 5, Appli
9	76.4	20.5	1500	17	US-10-191-803-73	Sequence 73, Appl
10	76.4	20.5	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
11	63.8	17.2	6128	9	US-09-880-107-2393	Sequence 2393, Ap
12	63.8	17.2	6128	22	US-10-756-149-1484	Sequence 1484, Ap
13	63.8	17.2	9173	22	US-10-893-315-126	Sequence 126, App
14	63.8	17.2	9174	22	US-10-893-315-160	Sequence 160, App
15	55.2	14.8	13011	17	US-10-388-934-36	Sequence 36, Appl
16	55.2	14.8	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
17	51	13.7	51	9	US-09-972-916A-1	Sequence 1, Appli
18	51	13.7	270	9	US-09-972-916A-3	Sequence 3, Appli
19	40.2	10.8	31703	17	US-10-085-117-172	Sequence 172, App
20	37.8	10.2	7061	15	US-10-311-455-970	Sequence 970, App
21	37.8	10.2	7061	17	US-10-221-613-148	Sequence 148, App
22	37.6	10.1	3191	16	US-10-029-386-20306	Sequence 20306, A
23	37.4	10.1	12850	14	US-10-017-122-1	Sequence 1, Appli
24	36.6	9.8	761	17	US-10-264-237-828	Sequence 828, App
25	36.6	9.8	1678	17	US-10-094-749-9	Sequence 9, Appli
26	36.2	9.7	586	16	US-10-029-386-5186	Sequence 5186, Ap
27	35.8	9.6	11009	9	US-09-845-583-1	Sequence 1, Appli
28	35.8	9.6	11009	14	US-10-037-182-3	Sequence 3, Appli
29	35.8	9.6	11009	21	US-10-764-420-947	Sequence 947, App
30	35.6	9.6	490	10	US-09-918-995-22070	Sequence 22070, A
31	35.4	9.5	1934	18	US-10-424-599-61023	Sequence 61023, A
32	35.2	9.5	1687	20	US-10-739-930-5461	Sequence 5461, Ap
33	35	9.4	769	13	US-10-027-632-164336	Sequence 164336,
34	35	9.4	769	17	US-10-027-632-164336	Sequence 164336,
35	35	9.4	3895	14	US-10-011-585A-76	Sequence 76, Appl
36	34.4	9.2	985	19	US-10-767-701-15088	Sequence 15088, A
37	34	9.1	10537	9	US-09-764-869-1269	Sequence 1269, Ap
38	34	9.1	10537	14	US-10-091-504-1269	Sequence 1269, Ap
39	34	9.1	10537	17	US-10-227-577-1269	Sequence 1270, Ap
40	34	9.1	10543	9	US-09-764-869-1270	Sequence 1270, Ap
41	34	9.1	10543	14	US-10-091-504-1270	Sequence 1270, Ap
42	34	9.1	10543	17	US-10-227-577-1270	Sequence 1270, Ap
43	34	9.1	49979	19	US-10-741-601-5746	Sequence 5746, Ap
44	34	9.1	49979	21	US-10-741-601-5746	Sequence 17905, A
45	33.4	9.0	201	20	US-10-719-993-35505	Sequence 35505, A

ALIGNMENTS

RESULT 1
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 100.0%; Score 372; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 2, 5e-116;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TACACTGGGGGCGGAGTCCAGGAACACGAGGAGTCCCGGCGCCCATTTACACTGGG 60
DB 1 TACACTGGGGGCGGAGTCCAGGAACACGAGGAGTCCCGGCGCCCATTTACACTGGG 60
QY 61 GGCCAGAGTCCAGGAACACGAGGAGTCCCGGCGCCCATTTACACTGGGCGGCGAGAGT 120

```
Db      61 GGCACAGATCCAGAAACAAGGAGTGCCTGGGCCCATGTACACTGGGGCCAGAGT 120
Qy      121 CAGAGAACCAAGGAGTGCCTGGGCCCATGTACAAACAACTATTGTA 180
Db      121 CCAGAACCAAGGAGTGCCTGGGCCCATGTACAAACAACTATTGTA 180
Qy      181 CAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGAGCCGCTT 240
Db      181 CAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGAGCCGCTT 240
Qy      241 CATTAAGCCCTGGTATGCGCAGCAGCATGTGCACTGCCGAGACAAACCCAG 300
Db      241 CATTAAGCCCTGGTATGCGCAGCAGCATGTGCACTGCCGAGACAAACCCAG 300
Qy      301 CAGACATTGAACACTGACACAGGCCCATCTGCCCCAGAGCTGTGACCACTTCGCTA 360
Db      301 CAGACATTGAACACTGACACAGGCCCATCTGCCCCAGAGCTGTGACCACTTCGCTA 360
Qy      361 CTAGCTAGCCGC 372
Db      361 CTAGCTAGCCGC 372
```

```
RESULT 2
US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4
```

Query Match 86.3%; Score 321; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 5.8e-99;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      52 TACACTGGGGGCAAGTCCAGAAACAAGGAGTGCCTGGGCCCATGTACACTGGG 111
Db      1 TACACTGGGGGCAAGTCCAGAAACAAGGAGTGCCTGGGCCCATGTACACTGGG 60
Qy      112 GGCAGAGTCCAGAAACAAGGAGTGCCTGGGCCCATGTACAAACAAACT 171
Db      61 GGCAGAGTCCAGAAACAAGGAGTGCCTGGGCCCATGTACAAACAAACT 120
Qy      172 TATTTTGAACAGGGGATCTAGACGCTGCGCTTACAAATTAACCCGCTGCGAG 231
Db      121 TATTTTGAACAGGGGATCTAGACGCTGCGCTTACAAATTAACCCGCTGCGAG 180
Qy      232 CCAGCCCTTCATTAAGGCTTGGGTATGCGCAGCAGATGTCCTGCGCGAGACA 291
Db      181 CCAGCCCTTCATTAAGGCTTGGGTATGCGCAGCAGATGTCCTGCGCGAGACA 240
Qy      292 CAACCCAGCAGCATTTGAACACTGACACAGGCCCATCTGCCCCAGAGCTGTGACCA 351
Db      241 CAACCCAGCAGCATTTGAACACTGACACAGGCCCATCTGCCCCAGAGCTGTGACCA 300
Qy      352 CTTCGCTACTAGTACGCGC 372
Db      301 CTTCGCTACTAGTACGCGC 321
```

```
RESULT 3
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6
```

Query Match 60.1%; Score 223.6; DB 9; Length 423;
Best Local Similarity 76.2%; Pred. No. 9.6e-66;
Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

```
Qy      2 ACACCTGGGGGCAAGTCCAGAAACAAGGAGTGCCTGGGCCCATGTACACTGGG 61
Db      34 ACTGTGCCCCCACTGTACATGAGGCGCAGCGGAGCATCCGCTGCTTCTGACTTGGCC 93
Qy      62 GCCAGATCCAGAAACAAGGAGTGCCTGGGCCCATGTACACTGGGGCCAGAGTC 121
Db      94 CCCAGTGTACATGAGGCGCAGCGGAGCATCCCGTGTTCCTGACTGTGCCCCAGAGTGA 153
Qy      122 CAGAACCAAGGAGTGCCTGGGCCCATG-----TCAGAGCA 162
Db      154 CATGGGCGACAGGGGCACTCCGCTGTCTGACTGTGCCCCCACTGTATCAAGCA 213
Qy      163 AAACAACTATTTTGAACAGGGGATCTAGACGCTGCGCTTACAAATTAACCGCT 222
Db      214 AAACAACTATTTTGAACAGGGGATCTAGACGCTGCGCTTACAAATTAACCGCT 273
Qy      223 GCTCCGAGCAGCCCTTCATTAAGGCTTGGGTATGCGCAGCAGATGTCCTGCTGCC 282
Db      274 GCTCCGAGCAGCCCTTCATTAAGGCTTGGGTATGCGCAGCAGATGTCCTGCTGCC 333
Qy      283 GCCAGACACAAACCCAGCAGCATTTGAACACTGACACAGGCCCATCTGCCCCAGAGCTG 342
Db      334 GCCAGACACAAACCCAGCAGCATTTGAACACTGACACAGGCCCATCTGCCCCAGAGCTG 393
Qy      343 TGACCACCACTTCGCTACTAGTACGCGC 372
Db      394 TGACCACCACTTCGCTACTAGTACGCGC 423
```

```
RESULT 4
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 59.0%; Score 219.4; DB 9; Length 270;
Best Local Similarity 99.5%; Pred. No. 2.3e-64;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 152 TGTCAAGCAAACTTTTGAACAGGGGATCTGACAGCGTGGCCCTGACAT 211
DB 50 TATCAAGCAAACTTTTGAACAGGGGATCTGACAGCGTGGCCCTGACAT 109
QY 212 CATTAACCGGTGCTGCGGAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGATG 271
DB 110 CATTAACCGGTGCTGCGGAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGATG 169
QY 272 GTTCACTGCGCGCGGAGACAAACCAGCAGCATTTGAACACTGCACAGGCCCATCTGC 331
DB 170 GTTCACTGCGCGCGGAGACAAACCAGCAGCATTTGAACACTGCACAGGCCCATCTGC 229
QY 332 CGAGAGCTGTGACCACTTCGGCTACTAGTACGCCG 372
DB 230 CGAGAGCTGTGACCACTTCGGCTACTAGTACGCCG 270

RESULT 5
US-09-972-916A-2

Sequence 2, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

CURRENT FILING DATE: 2001-10-10

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 2

LENGTH: 219

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

US-09-972-916A-2

Query Match 58.9%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 3e-64;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TCACAGCAAACTTTTGAACAGGGGATCTGACAGCGTGGCCCTGACATCA 213
DB 1 TCACAGCAAACTTTTGAACAGGGGATCTGACAGCGTGGCCCTGACATCA 60
QY 214 TTAACCGGTGCTGCGGAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGATGT 273
DB 61 TTAACCGGTGCTGCGGAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGATGT 120
QY 274 CCAGTCCCGCGGAGACAAACCAGCAGCATTTGAACACTGCACAGGCCCATCTGCC 333
DB 121 CCAGTCCCGCGGAGACAAACCAGCAGCATTTGAACACTGCACAGGCCCATCTGCC 180
QY 334 AGAGAGCTGTGACCACTTCGGCTACTAGTACGCCG 372
DB 181 AGAGAGCTGTGACCACTTCGGCTACTAGTACGCCG 219

RESULT 6
US-09-972-916A-6/c

Sequence 6, Application US/09972916A

Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

CURRENT FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 6

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized

US-09-972-916A-6

Query Match 41.4%; Score 154; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 5e-42;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTGGGGGCGAGTTCAGAAACAGGGAGTCCCGTGGCCCATGTACATGGGCGGAGT 60
DB 204 TAACTGGGGGCGAGTTCAGAAACAGGGAGTCCCGTGGCCCATGTACATGGGCGGAGT 145
QY 61 GGCAGAGTCCAGGAACAGGGAGTCCCGTGGCCCATGTACATGGGCGGAGT 120
DB 144 GGCAGAGTCCAGGAACAGGGAGTCCCGTGGCCCATGTACATGGGCGGAGT 85
QY 121 CCAGAACACAGGAGTCCCGTGGCCCATGT 154
DB 84 CCAGAACACAGGAGTCCCGTGGCCCATGT 51

RESULT 7
US-09-917-800A-1608

Sequence 1608, Application US/09917800A
Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Caselle, Arthur

APPLICANT: Elshoff, Michael

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5018-US

CURRENT APPLICATION NUMBER: US/09/917,800A

CURRENT FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303,459

PRIOR FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1608

LENGTH: 1500

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144

US-09-917-800A-1608

PRIOR FILING DATE: 2002-05-06

Query Match	20.5%	Score 76.4;	DB 18;	Length 1500;
Best Local Similarity	98.7%	Pred. No. 1.9e-15;		
Matches 77; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 117.486 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372

Sequence: 1 taccatggggggccagatgcc.....ttccgctactagctagccgc 372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/pdata/1/ina/5A_COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/pdata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63.8	17.2	9173	US-09-949-001-30	Sequence 30, Appl
2	55.2	14.8	9174	US-09-949-001-36	Sequence 36, Appl
3	55.2	14.8	194	US-08-945-140-1	Sequence 1, Appl
4	55.2	14.8	13011	US-08-791-849A-14	Sequence 14, Appl
5	40.2	10.8	15108	US-09-949-016-11786	Sequence 11786, A
6	40.2	10.8	15108	US-09-949-016-117205	Sequence 11705, A
7	36.6	9.8	3054	US-09-484-970B-138	Sequence 138, App
8	36.2	9.7	35471	US-09-949-016-12957	Sequence 12957, A
9	36	9.7	40586	US-09-949-016-16965	Sequence 16965, A
10	35.8	9.6	11009	US-09-845-583A-1	Sequence 1, Appl
11	35.8	9.6	16520	US-09-949-016-11710	Sequence 14710, A
12	35.8	9.6	16520	US-09-949-016-15394	Sequence 15394, A
13	35.6	9.6	76767	US-09-949-016-12147	Sequence 12147, A
14	35.6	9.6	76767	US-09-949-016-17361	Sequence 17361, A
15	35	9.4	601	US-09-949-016-19194	Sequence 19194, A
16	35	9.4	601	US-09-949-016-73784	Sequence 73784, A
17	35	9.4	2293	US-09-949-016-2153	Sequence 2153, Ap
18	35	9.4	2408	US-09-949-016-74	Sequence 74, Appl
19	35	9.4	7561	US-09-949-016-11816	Sequence 11816, A
20	35	9.4	7562	US-09-949-016-13895	Sequence 13895, A
21	34	9.1	30244	US-09-949-016-13208	Sequence 13208, A
22	34	9.1	30245	US-09-949-016-13550	Sequence 13550, A
23	33.4	9.0	364	US-09-270-767-28006	Sequence 28006, A
24	33.4	9.0	601	US-09-949-016-35649	Sequence 35649, A
25	33.4	9.0	601	US-09-949-016-177807	Sequence 177807, A
26	33.4	9.0	761	US-09-270-767-13109	Sequence 13109, A
27	32.8	8.8	749	US-09-257-583-12	Sequence 12, Appl

c 28	32.8	8.8	2130	4	US-09-909-962A-6	Sequence 6, Appl
c 29	32.8	8.8	2130	4	US-09-909-962A-7	Sequence 7, Appl
c 30	32.6	8.8	907	4	US-09-585-645A-55	Sequence 65, Appl
c 31	32.6	8.8	134008	4	US-09-949-016-11841	Sequence 11841, A
c 32	32.2	8.7	505	4	US-09-621-976-15639	Sequence 15639, A
c 33	32.2	8.7	601	4	US-09-949-016-35650	Sequence 35650, A
c 34	32.2	8.7	601	4	US-09-949-016-35651	Sequence 35651, A
c 35	32.2	8.7	601	4	US-09-949-016-177808	Sequence 177808, A
c 36	32.2	8.7	601	4	US-09-949-016-177809	Sequence 177809, A
c 37	32.2	8.7	1281	4	US-09-902-540-3669	Sequence 3669, Ap
c 38	32.2	8.7	4447	2	US-08-304-309-3	Sequence 3, Appl
c 39	32.2	8.7	4447	3	US-08-991-942-3	Sequence 1177, Ap
c 40	32.2	8.7	23847	4	US-09-902-540-1177	Sequence 16824, A
c 41	32.2	8.7	46343	4	US-09-949-016-16824	Sequence 11745, A
c 42	32	8.6	36180	4	US-09-949-016-11745	Sequence 16163, A
c 43	32	8.6	36181	4	US-09-949-016-16163	Sequence 10411, A
c 44	31.8	8.5	2194	4	US-09-270-767-10411	Sequence 404, App
c 45	31.6	8.5	1348	4	US-09-774-528-404	

ALIGNMENTS

```

RESULT 1
US-09-949-001-30
; Sequence 30, Application US/09949001
; Patent No. 6623336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FASTSQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-30

Query Match      17.2%; Score 63.8; DB 4; Length 9173;
Best Local Similarity 63.2%; Pred. No. 2.2e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCACGAAACCAAGGAGTCCCGCGCCGATGTCACAAACAAACTTATTGGA 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1842 TCCTCCACCAAGCGGTTGCTGAGGCTTGGGTGCTACCAAACTTATTGGA 1901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 ACACGGGAGTCTAGACAGCTGCTGCAATCAATTAATTAATTAATTAATTA 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1902 ACATCAGCTCTGAGCGTGGCGGCTGCAATCAATTAATTAATTAATTAATTA 1961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 AGCAGCCTTCATPAGGCTCTGGTATGCGCAGCAGCATGCTTCACTGCGCCGAGA 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1962 CCTGCGCTTATPAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 CACAAACCAAGCAGCATGTAACCTGC-ACACGCGCATGCGCAGAGACTGTGACCA 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2018 ---CATCAGCAGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 CCACTTC 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2074 CCGCTCC 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-949-001-36
; Sequence 36, Application US/09949001
; Patent No. 6625336

```

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match 17.2%; Score 63.8; DB 4; Length 9174;
Best Local Similarity 63.2%; Pred. No. 2.2e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCCAGAACACGGAGTGCCTCCGCGCCATGTCACAGCAAAACAATTATTGGA 179
DB 1842 TCCTCCACACGCGGTTGCGAGGAGCTTGAGTGCACAGCAAAACAATTATTGGA 1901
QY 180 ACACGGGATCTTACACAGCTGCTGACATCATTAACCC-----GTGCTGCG 229
DB 1902 ACACCTACGCTCTTACAGTGCAGCGCTGCAATCATTAACCTCTGTCAGATGCGCGG 1961
QY 230 AGCAGCCCTTCATTAAGCCCTGAGTATGCGCAGCAGATGTCACCTGCGCGCGAGA 289
DB 1962 CCTGTCCTTATTAAGTGCAGCGCTGTCAGAGAGATGCGCACCGCATCC---- 2017
QY 290 CACAACCCAGAGACATTTGAACACTGC-ACA CGCCATCTGCCAGAGAGCTGAGCA 348
DB 2018 ---CATCAGAGACATCTGCGCGCGCGCGCCACCTCCAGAGAGACATGCGCA 2073
QY 349 CCACTTC 355
DB 2074 CCGCTCC 2080

RESULT 3
US-08-945-140-1/c
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Bq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match 14.8%; Score 55.2; DB 3; Length 194;
Best Local Similarity 88.2%; Pred. No. 2.6e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCCGTGCGCCCATGATGACCTGAGGCGCAGAGTCCAGAACACGGAAGTCCCGGTG 94
DB 78 TGCCACGGAAGCCTGTGACACTGGGCGCCAGAGTCCAGAACACGGAAGTCCCGGTG 19
QY 95 GCGCCATG 102
DB 18 GCGCCATG 11

RESULT 4
US-08-791-849A-14/c
Sequence 14, Application US/08791849A
Patent No. 591449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
TITLE OF INVENTION: Lipid Content in Plant Seed
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
LOCATION: 9480..10162)
US-08-791-849A-14

Query Match 14.8%; Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.7e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 GTGCCCCGTGCGCCATGTACACTGGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 94
DB 3088 TGCACACGAGACCTGTATGACTGGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 3029
QY 95 CGCCCATG 102
DB 3028 CGCCCATG 3021

RESULT 5
US-09-949-016-11786/c
Sequence 11786, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11786
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-11786

Query Match 10.8%; Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 34 GTGCCCCGTGCGCCATGTACACTGGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 93
DB 2417 GTTCATTGGAAGCCCTGTATGCTCAGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357

RESULT 6
US-09-949-016-17205/c

Sequence 17205, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17205
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-17205

Query Match 10.8%; Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 34 GTGCCCCGTGCGCCATGTACACTGGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 93
DB 2417 GTTCATTGGAAGCCCTGTATGCTCAGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357

RESULT 7
US-09-484-970B-138/c
Sequence 138, Application US/09484970B
Patent No. 6426186

GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmutch, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 138
LENGTH: 3054
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6426186 336987.1CB1
US-09-484-970B-138

Query Match 9.8%; Score 36.6; DB 3; Length 3054;
Best Local Similarity 62.6%; Pred. No. 0.11;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 ACACGTGGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGTGCCTCATGTACACTGGGG 61
DB 355 ACTTCGGAAGCAAGCCCTGTATGCTCAGGGGCGCAGATCCAGAACCAACGGAGTGCCTCGT 296
QY 62 GCCAGATCCAGAACCAACGGAGTGCCTCGG 92
DB 295 GCTGAGGCTTGGAGCGAAGGAGAGCCCGG 265

RESULT 8

US-09-949-016-14710

Query Match	9.6%	Score 35.8;	DB 4;	Length 16520;
Best Local Similarity	52.3%	Pred. No. 0.43;		
Matches 79;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0

QY	7	GGGGGGCCAGAGTCCAGAAACACGGGAGTGGCCCGTGGCCCATGTATCACTGGGGGCGAC	66
Db	3583	GGGTGTGGGTGTCCCGGAGTGTGGGTGTCCCGGGGGCGTGGGTGTCCCGGAGTGTGGG	3642
QY	67	AGTCCAGGAACACGGGAGTGGCCCCGTGGCCCATGTATCACTGGGGGCGAGAGTCCAGGA	126
Db	3643	TGTCCCCGGGGGCGGTGGGTGTCCCGGAGTGTGGGTGTCCCGGGGCGTGGGTGTCCCGGG	3702
QY	127	ACCACGGAGATGCCCCCTGGCGCCATGTCTAC	157
Db	3703	AGTGTGGGTGTCCCGGGGCTGTGGGTGTCTCC	3733

RESULT 12

US-09-949-016-15394
 : Sequence 15394, Application US/09949016
 : Patent No. 6812339
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig et al.
 : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 : WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: CL001307
 : CURRENT APPLICATION NUMBER: US/09/949, 016
 : CURRENT FILING DATE: 2000-04-14
 : PRIOR APPLICATION NUMBER: 60/241,755
 : PRIOR FILING DATE: 2000-10-20
 : PRIOR APPLICATION NUMBER: 60/237,768
 : PRIOR FILING DATE: 2000-10-03
 : PRIOR APPLICATION NUMBER: 60/231,498
 : PRIOR FILING DATE: 2000-09-08
 : NUMBER OF SEQ ID NOS: 207012
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 15394
 : LENGTH: 16520
 : TYPE: DNA
 : ORGANISM: Human
 : US-09-949-016-15394

Query Match	9.6%	Score 35.8;	DB 4;	Length 16520;
Best Local Similarity	52.3%	Pred. No. 0.43;		
Matches	79;	Conservative	0;	Mismatches 72;
			Indels	0;
			Gaps	0

QY	7	GGGGGCGAGATCTTCAGGAAACACGGGAAATGCCCGCTGGCCCATTTACACTGGGGGCGAC	66
Db	3583	GGGTGTGGGTGTCTCCGGAGTGTGGGTGTCTCCGGGGCGTGGGTGTCTCCGGAGTGTGGG	3644
QY	67	AGTCGAGAAACACGGGAGTGTCCCGTGCGCCCATGTAACTGGGGGCGCAGATCCAGGA	126
Db	3643	TGTCCCGGGGGGTGGGTGTCTCCGGGAAATGTGGGTGTCTCCGGGGCGTGGGTGTCTCCGGG	3702
QY	127	ACCAACGGGAGTGTCCCGTGTGGCGCCCATGTTCAC	157
Db	3703	AGTGTGGGTGTCTCCGGGGGTGTGGGTGTCTCC	3733

RESULT 13

US-09-949-016-12147/C
Sequence 12147, Application US/09949016
Parent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL0011307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

```

1 PRIOR FILING DATE 2000-10-20
2 PRIOR APPLICATION NUMBER: 60/237,768
3 PRIOR FILING DATE: 2000-10-03
4 PRIOR APPLICATION NUMBER: 60/231,498
5 PRIOR FILING DATE: 2000-09-06
6 NUMBER OF SEQ ID NOS: 207012
7 SOFTWARE: FASTQ for Windows Version 4.0.

```

Query Match	9.6%	Score 35.6;	DB 4;	Length 767677;
Best Local Similarity	51.9%	Pred. No. 2.8;		
Matches 80;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0

QY	7	GGGGGGCAGAGTCTCAGGAAACACGGGAAATGGCCCGCTGGCCCAATGTACACTGGGGGGCAG	66
Db	562461	GGGGCCCTTGAAGAAAGAAATCAGGGAAATCCCCGACCCACCAAGTCCACACGAGAGCCCTG	5624202
QY	67	AGTCCAGAAACACGGGAGTGCCTCCCGTCGCGCCATGTACACTGGGGGCCAGATCCAGGA	126
Db	562401	AGTGTGAATATACGGGAACTCCTCTACACCTTACAGTACTCACTCAGGGGCCCTTGAAGAGAG	5623422
QY	127	ACCCAGGAGAGTCCCGCTGGCGCCCATGTCAACAG	160
Db	562341	ATCACGGGACTCCTCGACCCACAGTGTCAACAG	562308

RESULT 14

```

US-09-949-016-17361/c      ; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

```

Query Match	9.6%	Score 35.6;	DB 4;	Length 767677;
Best Local Similarity	51.9%;	Pred. No. 2.8;		
Matches 80;	Conservative 0;	Mismatches 74;	Indels 0;	Gaps 0;

Qy 7 GGGGGCCGAGATCCAGGAACACGGGAGTGGCCCGTGGCCATGTACATGGGGGCCAG 66
Db 562461 GGGGGCCCTGAGCAAGAGATCACGGGACTCCCGCACCCACAGTCCACAGGAGCCCTG 562402
Qy 67 AGTCCAGGAACACGGGAGTGGCCCGTGGCCATGTACATGGGGGCCAGAGATCCAGGA 126

Db 562401 AGTGTGAGATACGGGACTCCCTACACCCCTACAGTACTACGGGGCCTTGAGCAAGAG 562342
 QY 127 ACCACGGAGTGCCCGGTGCGCCCATGTCAAG 160
 Db 562341 ATCACGGGACTCCCTTGACACCCACAGTCCACAG 562308

RESULT 15
 US-09-949-016-19194/c
 ; Sequence 19194, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 19194
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-19194

Query Match 9.4%; Score 35; DB 4; Length 601;
 Best Local Similarity 55.3%; Pred. No. 0.18; Mismatches 55; Indels 0; Gaps 0;
 Matches 68; Conservative 0;

QY 232 CCAGCCCTTCATTAAGGCCCTGGGTATGGCCAGCCAGCATGTCTCACTGCGCCCGAGACA 291
 Db 478 CCAGGCTTCTCGGGCGCTGTTCCTCTTCATCCAGCTTAGCCACCTGCGAGTTAGGCC 419
 QY 292 CAAACCCAGCGGACATTGAACCTGACACGGGCATCTGCCAGAGAGCTGTGACCACCA 351
 Db 418 CCCATCCAGAGACACATCCACCGGATCACAGCCCTGCCACACCATCTCATGATAGCC 359
 QY 352 CTT 354
 Db 358 CTT 356

Search completed: September 1, 2005, 07:07:16
 Job time : 123.652 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 31, 2005, 13:10:38 ; Search time 2633.79 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-6

Perfect score: 423
Sequence: 1 celtggggcgcagcgggcagc.....ttccgctaactagtagccgc 423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl1:
1: gb_bai:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	212.4	50.2	1181	10	RAT1IGFZ
2	212.4	50.2	185148	2	AC136382
3	188.4	44.5	1363	10	MMILGP
4	188.4	44.5	192843	10	AL607124
5	98.2	23.2	1569	10	BC013345
6	76.6	18.1	1507	10	AY560836
7	76.4	18.1	1500	6	AX401932
8	76.4	18.1	1500	6	AX827271
9	76.4	18.1	1500	10	RAT1IGFB
10	76.4	18.1	5001	6	AX163782
11	76.4	18.1	5001	10	RAT1IGFBA
12	65.4	15.5	1510	10	BC078889
13	62.6	14.8	3886	9	AY095345
14	60.8	14.4	448	11	G671393
15	60.8	14.4	6128	6	AX409747
16	60.8	14.4	6128	9	HUMIGFBP1A
17	60.8	14.4	6128	11	G19994
18	60.8	14.4	6480	9	HUMIGFBP1
19	60.8	14.4	9082	9	AY434089

20	60.8	14.4	69887	9	AC091524
21	59.2	14.0	141539	2	AC146152
22	59.2	14.0	189932	2	AC146117
23	59.2	14.0	200935	2	AC148834
24	55.8	13.2	125020	2	AF429315
25	55.8	13.2	256781	2	AC097952
26	55.8	13.2	259329	2	AC109570
27	55.2	13.0	194	6	AS7715
28	55.2	13.0	194	6	AR175909
29	55.2	13.0	13011	6	E14395
30	55.2	13.0	13011	6	AX827302
31	55.2	13.0	13011	10	RMLPKG
32	53.6	12.7	81704	9	AC110299
33	53.6	12.7	151700	9	AC133528
34	53.4	12.6	200340	9	AC100797
35	52.2	12.3	170654	10	AC116502
36	52.2	12.3	187397	2	AC114566
37	52.2	12.3	218657	10	AC110517
38	51.6	12.2	184889	9	AL442125
39	50.6	12.0	994	11	BY006814
40	50.6	12.0	107103	9	AL589702
41	50.6	12.0	145540	9	AP001052
42	50.6	12.0	184104	9	AC020709
43	50.6	12.0	340000	9	AP001752
44	50.4	11.9	895	9	HUMIGFBP1
45	50	11.8	53370	9	AL592071

ALIGNMENTS

RESULT 1	RAT1IGFZ	1181 bp	DNA	linear	ROD 27-APR-1993
LOCUS	Rattus norvegicus insulin-like growth factor gene fragment.				
DEFINITION	M84484.1 GI:204927				
ACCESSION	insulin-like growth factor.				
VERSION	Rattus norvegicus (Norway rat)				
KEYWORDS	Rattus norvegicus				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae;				
ORGANISM	Rattus				
REFERENCE	1 (bases 1 to 1181)				
AUTHORS	Uncerman,T.G., Jacson,R.G., McGary,E., Whalen,C. and Goswami,R.G.				
JOURNAL	Biochem. Biophys. Res. Commun. (1991) In press				
COMMENT	Original source text: Rattus norvegicus (strain Sprague-Dawley) male adult liver DNA.				
FEATURES	Location/Qualifiers				
source	1..1181				
	/organism="Rattus norvegicus"				
	/mol_type="genomic DNA"				
	/strain="Sprague-Dawley"				
	/db_xref="taxon:10116"				
	/sex="male"				
	/tissue_type="liver"				
	/dev_stage="adult"				
ORIGIN					
Query Match	50.2% Score 212.4; DB 10; Length 1181;				
Beet Local Similarity	99.5%; Pred. No. 1.2e-48; Indels 0; Gaps 0;				
Matches	213; Conservative 0; Mismatches 1;				
QY	205 TCACAGCAAAACAACTATTATTTGAACACGGGATCTTACACGCTCCCTGCACATCA	264			
DB	823 TCACAGCAAAACAACTATTATTTGAACACGGGATCTTACACGCTCCCTGCACATCA	882			
QY	265 TTAACCCGTGCTGCCAGCCAGCCCTTATTAAGGCGGTATGCGCCAGCCAGCATGCT	324			
DB	883 TTAACCCGTGCTGCCAGCCAGCCCTTATTAAGGCGGTATGCGCCAGCCAGCATGCT	942			
QY	325 CCACGCGCCGCGAGACCAAAACCGAGCATTTGAACACTGCACACGGCATCTGCGC	384			

Db 943 CCATGCCCCGAGAGACACAAACCCAGCAGATTGACACTGCAACGCGCATCTGCCC 1002
OY 385 AGAGAGCTGTGACACCACTTCGGCTACTAGCTA 418
Db 1003 AAGAGACTGTGACCACTTCGGCTACTACTTA 1036

RESULT 2
AC136382
LOCUS
DEFINITION Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
AC136382 185148 bp DNA linear HTG 01-NOV-2002
*** 63 unordered pieces.
AC136382
VERSION AC136382.1 GI:24462257
KEYWORDS HTG; HTGS PHASE1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 185148)
Muzny,D,Marie, Metzger,M, Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
David, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, A., Guerra, W., Guevara, M.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kows, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshaw, L., Louie, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M., McNeill, T., Meenan, E., Milosavljevic, A.,
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G.,
Olanunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A.,
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojars, A.,
Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C.D., Smaaj, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseana, D.,
Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, F., White, F., Williams, G., Willson, R., Wleczek, R.,
Wooden, H., Wooley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R.,
Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
Rat Genome Sequencing Consortium.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (01-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KOGS
Center clone name: CH230-97018
----- Summary Statistics
Sequencing vector: plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 135613 bases at least Q40
Consensus quality: 140849 bases at least Q30
Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1011 11010: contig of 1010 bp in length
1111 2608: contig of 1498 bp in length
2609 2708: gap of unknown length
2709 3915: contig of 1207 bp in length
3916 4015: gap of unknown length
4016 5547: contig of 1532 bp in length
5548 5648: gap of unknown length
5648 7120: contig of 1472 bp in length
7120 7220: gap of unknown length
7220 8885: contig of 1666 bp in length
8885 8985: gap of unknown length
8985 10258: contig of 1273 bp in length
10259 10358: gap of unknown length
10359 11626: contig of 1268 bp in length
11627 11726: gap of unknown length
11727 13488: contig of 1762 bp in length
13489 13588: gap of unknown length
13589 15122: contig of 1534 bp in length
15123 15222: gap of unknown length
15223 16946: contig of 1724 bp in length
16947 17047: gap of unknown length
17047 18164: contig of 1118 bp in length
18165 18264: gap of unknown length
18265 19678: contig of 1414 bp in length
19679 21687: gap of unknown length
21688 21877: contig of 1909 bp in length
21878 22892: gap of unknown length
22893 22992: contig of 1105 bp in length
22993 24336: gap of unknown length
24337 24437: contig of 1344 bp in length
24438 26117: gap of unknown length
26118 26577: contig of 2181 bp in length
26578 26718: gap of unknown length
26718 28577: contig of 1860 bp in length
28578 30823: gap of unknown length
30824 30923: contig of 2146 bp in length
30924 32266: gap of unknown length
32267 32366: contig of 1343 bp in length
32367 33894: gap of unknown length
33895 33994: contig of 1528 bp in length
33995 35373: gap of unknown length
35374 35373: contig of 1379 bp in length

35374 35473: gap of unknown length
35474 37295: contig of 1822 bp in length
37296 37395: gap of unknown length
37396 38420: contig of 1025 bp in length
38421 38520: gap of unknown length
38521 39974: contig of 1454 bp in length
39975 40074: gap of unknown length
40075 42283: contig of 2209 bp in length
42284 42383: gap of unknown length
42384 43721: contig of 1338 bp in length
43722 43821: gap of unknown length
43822 46538: contig of 2717 bp in length
46539 46638: gap of unknown length
46639 48621: contig of 1993 bp in length
48622 48721: gap of unknown length
48722 50602: contig of 1881 bp in length
50603 50702: gap of unknown length
50703 53841: contig of 3139 bp in length
53842 53941: gap of unknown length
53942 56195: contig of 2254 bp in length
56196 56295: gap of unknown length
56296 58949: contig of 2654 bp in length
58950 61032: contig of 1983 bp in length
61033 61132: gap of unknown length
61133 62942: contig of 1810 bp in length
62943 63042: gap of unknown length
63043 66045: contig of 3003 bp in length
66046 66145: gap of unknown length
66146 68197: contig of 2052 bp in length
68198 68297: gap of unknown length
68298 72235: contig of 3938 bp in length
72236 72335: gap of unknown length
72336 75903: contig of 3468 bp in length
75904 78355: contig of 2452 bp in length
78356 78455: gap of unknown length
78456 81458: contig of 3003 bp in length
81459 81558: gap of unknown length
81559 84673: contig of 3115 bp in length
84674 87739: gap of unknown length
87740 87739: contig of 2966 bp in length
87740 87839: gap of unknown length
87840 91535: contig of 3696 bp in length
91536 91635: gap of unknown length
91636 95522: contig of 3887 bp in length
95523 95622: gap of unknown length
95623 98730: contig of 3108 bp in length
98731 98830: gap of unknown length
98831 101618: contig of 2788 bp in length
101619 101718: gap of unknown length
101719 105535: contig of 3817 bp in length
105536 105635: gap of unknown length
105636 109393: contig of 3758 bp in length
109394 109493: gap of unknown length
109494 113664: contig of 4071 bp in length
113665 113664: gap of unknown length
113665 117630: contig of 3966 bp in length
117631 117730: gap of unknown length
117731 122502: contig of 4772 bp in length
122503 122602: gap of unknown length
122603 126644: contig of 4042 bp in length
126645 126744: gap of unknown length

Query Match 50.2%; Score 212.4; DB 2; Length 185148;
Beet Local Similarity 99.5%; Pred. No. 1,le=48;
Matches 213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 205 TCACAAAGCAAAACAACTTATTGTAACACGGGAGATCTGAGACGCTGCGCTGACAAATCA 264
DB 99253 TCACAAAGCAAAACAACTTATTGTAACACGGGAGATCTGAGACGCTGCGCTGACAAATCA 99312
QY 265 TTAACCGTGTGCGAGCGACGCCCTTCAATTAAGCGCCCTGGGTATGCGCCAGCAGATGCT 324

DB 99313 TTAACCGTGTGCGAGCGACGCCCTTCAATTAAGCGCCCTGGGTATGCGCCAGCAGATGCT 99372
QY 325 CCATGCCCCCGGAGACAAACCCAGGAGACATTGAACACTGACAGCGCCATCTGCC 384
DB 99373 CCATGCCCCCGGAGACAAACCCAGGAGACATTGAACACTGACAGCGCCATCTGCC 99432
QY 385 AGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 418
DB 99433 AGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 99466
RESULT 3
LOCUS MMTLG 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493
VERSION X67493.1 GI:52659
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1363)
Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
Hepatology 19 (3), 656-665 (1994)
JOURNAL MEDLINE 94164648
PUBMED 7509771
REFERENCE
2 (bases 1 to 1363)
Mohn,K.L., Waddie,J.R. and Taub,R.
Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
Nucleic Acids Res.
REFERENCE
3 (bases 1 to 1363)
Taub,R.A.
Direct Submission
Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source location/Qualifiers
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="UM2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
792..1336
/number=1
792
/note="mRNA cap site"
964..1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

ORIGIN

/translation="MPBFLTVSWPFLILISFOIGVANAQPWMCAPCTAERLGLCP
PVPASCPBISRPAGCCGCTCALPMGAAAGVATARCACAGLSICRALPGEPRPLIALTRG
QGASLPEPAPATSTLFSQHE"

Query Match 44.5%; Score 188.4; DB 10; Length 1363;
Best Local Similarity 92.5%; Pred. No. 6.1e-42;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 205 TCACAGCAAAACAACTATTATTGAAACAGGGGATCTTAGACAGCGTCCCTGCAATCA 264
|||||
DB 677 TCACAGCAAAACAACTATTATTGAAACAGGGGATCTTAGACAGCGTCCCTGCAATCA 736
|||||

QY 265 TTAAACCGTCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGTATGCGCCAGCAGCATGCT 324
|||||
DB 737 TTAACTGTGCGGAGCCAGCCGCTTCATTAAGGCGCTGGGTATGCGCCAGCAGCATGCT 796
|||||

QY 325 CCACCTGCGCGGAGGAGACAAACCCAGGAGCATTTGAACATGTGACACAGGCGCATCTGCC 384
|||||
DB 797 CCACCTGCGCGGAGGAGACAAACCCAGGAGCATTTGAACATGTGACACAGGCGCATCTGCC 856
|||||

QY 385 AGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 418
|||||
DB 857 AGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 890
|||||

RESULT 4
AL607124
LOCUS Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
DEFINITION sequence.
ACCESSION AL607124
VERSION AL607124.15 GI:20145926
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Oliver K.
TITLE Direct Submission
JOURNAL Submitted (11-Apr-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humgeny@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES

source
1..192843
location/Oallifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

ORIGIN

/clone="RP23-20C9"
/clone_1id="RPCI-23"

Query Match 44.5%; Score 188.4; DB 10; Length 192843;
Best Local Similarity 92.5%; Pred. No. 5.7e-42;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 205 TCACAGCAAAACAACTATTATTGAAACAGGGGATCTTAGACAGCGTCCCTGCAATCA 264
|||||
DB 2429 TCACAGCAAAACAACTATTATTGAAACAGGGGATCTTAGACAGCGTCCCTGCAATCA 2488
|||||

QY 265 TTAAACCGTCTGCGGAGCCAGCCCTTCATTAAGGCGCTGGGTATGCGCCAGCAGCATGCT 324
|||||
DB 2489 TTAACTGTGCGGAGCCAGCCGCTTCATTAAGGCGCTGGGTATGCGCCAGCAGCATGCT 2548
|||||

QY 325 CCACCTGCGCGGAGGAGACAAACCCAGGAGCATTTGAACATGTGACACAGGCGCATCTGCC 384
|||||
DB 2549 CCACCTGCGCGGAGGAGACAAACCCAGGAGCATTTGAACATGTGACACAGGCGCATCTGCC 2608
|||||

QY 385 AGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 418
|||||
DB 2609 AGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 2642
|||||

RESULT 5
BC013345
LOCUS Mus musculus insulin-like growth factor binding protein 1, mRNA
DEFINITION (CDNA clone MGC:114075 IMAGE:4161889), complete cds.
ACCESSION BC013345
VERSION BC013345.1 GI:15426482
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,
Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
Dietcheno L., Marusina K., Farmer A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Scheer T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McEwen P.J.,
McKernan K.U., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Young A.C., Shvachenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalski U., Smallos D.E.,
Scherer A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1569)

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL
DIRECT SUBMISSION
Submitted (31-Aug-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: ggapds-remail.nih.gov
Tissue Procurement: Jeffrey B. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIN)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN.ac: http://image.llnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..1569
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:14075 IMAGE:4161889"
/issue_type="liver, normal. 5 month old male mouse."
/clone_id="NCI CGAP_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocustID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH13345.1"
/db_xref="GI:15426483"
/db_xref="MGI:96436"
/db_xref="LocustID:16006"
/translation="MPEPLTVSWPILLSPQIVAAAGAPQPHHCAPCTARLALCP
PVPASCELSRPAAGCCCTCALPMAAGCVATNACAGLSCTRALPGEPRLHALTRG
OGACVPEPAAPSTLSSQHEAKAAVVASADELSPEPMTESQLDSHLMASPERD
OPIIMNAISTYSMRAREIADIKWKEPCRELKYLRLAAQKAGBEIYKFLPN
CNKGFYHSKQETSIDSEALCMCVPMGSRIPGSLERDPPCHQYFNVQN"

CDS

23.2%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 1.2e-16;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

ORIGIN

Query Match 23.2%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 1.2e-16;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 308 TGGCCAGCAGAGTGTCCAGTCCGCGGAGAGCAACAACGAGAGCATTTGAACACTG 367
DB 20 TGGGCGAGCAGATGTCCAGTCCGCGGAGAGCAACAACGAGAGCATTTGAACACTG 79
QY 368 CACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 418
DB 80 CACAGCGCGCTCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 130

RESULT 6
LOCUS AY560836 1507 bp mRNA linear ROD 22-MAR-2004
DEFINITION *Spermophilus tridecemlineatus* insulin-like growth factor binding protein 1 (IGFBP1) mRNA, complete cds.
ACCESSION AY560836
VERSION AY560836.1 GI:45505308
KEYWORDS
SOURCE *Spermophilus tridecemlineatus* (thirteen-lined ground squirrel)
ORGANISM *Spermophilus tridecemlineatus*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae; Spermophilus.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Li,Y., Klimanis,D. and Hallenbeck,J.M.
TITLE Cloning and characterization of insulin-like growth factor binding protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL

Unpublished
2 (bases 1 to 1507)

REFERENCE Li,Y., Klimanis,D. and Hallenbeck,J.M.
AUTHORS Direct Submission
TITLE Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr., Bethesda, MD 20892, USA
JOURNAL

FEATURES

Location/Qualifiers

1..1507
/organism="Spermophilus tridecemlineatus"
/mol_type="mRNA"
/db_xref="taxon:43179"
1..1507
/gene="IGFBP1"
689..1507
/gene="IGFBP1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AA567029.1"
/db_xref="GI:45505309"
/translation="MPEVPAALMPFLLLVAVYSTVASSTOPPHHCAPCSAKALCP
FVPSCELSRPAAGCCCTCALPMAAGCVATNACAGLSCTRALPGEPRLHALTRG
OGACVPEPAAPSTLSSQHEAKAAVVASADELSPEPMTESQLDSHLMASPERD
OPIIMNAISTYSMRAREIADIKWKEPCRELKYLRLAAQKAGBEIYKFLPN
CNKGFYHSKQETSIDSEALCMCVPMGSRIPGSLERDPPCHQYFNVQN"

CDS

ORIGIN

Query Match 18.1%; Score 76.6; DB 10; Length 1507;
Best Local Similarity 67.6%; Pred. No. 1.3e-10;
Matches 144; Conservative 0; Mismatches 54; Indels 15; Gaps 2;

QY 204 ATCCAAAGCAAAACAACCTTTTGAACAGAGGATCTTACAGAGCCCTGACCAATC 263
DB 429 AGCAGAGCAAAACAACCTTTTGAACAGAGGATCTTACAGAGCCCTGACCAATC 488
QY 264 ATTA-----CCGCTGCGGAGCCAGCCCTCATAGGCGCTGAGTACGCA 313
DB 489 ATTAACCTTTAGTCCAGATGAGTGAACCGCGCCCTTTAAGGCAAGGCTGACCA 548
QY 314 GCCAGCATGTGTCACCTGCGCGCCGAGACAAACCCAGCAGCATTTGAACCTGACACG 373
DB 549 GCAAGCATGTGTCACCTGCGCGCCGAGACAAACCCAGCAGCATTTGAACCTTGA-----GCCACT 603
QY 374 GCCATGTGCCAGAGAGCTGTGACCAACCACTTC 406
DB 604 GCCAGCTGCCAGAGAGCATTTGACCACTGTCC 636

RESULT 7
LOCUS AX401932 1500 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1608 from Patent WO210453.
ACCESSION AX401932
VERSION AX401932.1 GI:21338112
KEYWORDS
SOURCE *Rattus norvegicus* (Norway rat)
ORGANISM *Rattus norvegicus*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1
AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Caetle,A.L. and Elashof,M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
source Location/Qualifiers

1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

Query Match	18.1%;	Score 76.4;	DB 6;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 1.5e-10;		
Matches 77;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
OY	341	CACAAACCCAGGACATTGAAACATGTGCACACGCGCATCTGCCAGAGAGCTGTGACAC	400	
Db	8	CACAAACCCAGGACATTGAAACATGTGCACACGCGCATCTGCCAGAGAGCTGTGACAC	67	
OY	401	CACCTTCGGCTACTACTA	418	
Db	68	CACCTTCGGCTACTACTA	85	
RESULT 8				
LOCUS	AX827271	1500 bp	DNA	linear
DEFINITION	Sequence 5 from Patent EP1344834.			
ACCESSION	AX827271			
VERSION	AX827271.1	GI:39837360		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Boess,F., Suter-Dick,L. and Wolf,D.			
TITLE	Methods for the toxicity prediction of a compound			
JOURNAL	Patent: EP 1344834-A 5 17-SEP-2003;			
FEATURES	F. HOFFMANN-LA ROCHE AG (CH)			
source	Location/Qualifiers			
ORIGIN	1..1500			
	/organism="Rattus norvegicus"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:10116"			
Query Match	18.1%;	Score 76.4;	DB 6;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 1.5e-10;		
Matches 77;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	341	CACAAACCCAGGACATTGAAACATGTGCACACGCGCATCTGCCAGAGAGCTGTGACAC	400	
Db	8	CACAAACCCAGGACATTGAAACATGTGCACACGCGCATCTGCCAGAGAGCTGTGACAC	67	
OY	401	CACCTTCGGCTACTACTA	418	
Db	68	CACCTTCGGCTACTACTA	85	
RESULT 9				
LOCUS	RAT1GFB	1500 bp	mRNA	linear
DEFINITION	Rat IGF binding protein-1 (rIFBp-1) mRNA, complete cds.			
ACCESSION	MS6634			
VERSION	MS6634.1	GI:204732		
KEYWORDS	IGF binding protein-1.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 1500)			
AUTHORS	Mohn,K.L., Melby,A.E., Tewari,D.S., Laz,T.M. and Taub,R.			
TITLE	The gene encoding rat insulinlike growth factor-binding protein 1			
JOURNAL	is rapidly and highly induced in regenerating liver			
MEDLINE	Mol. Cell. Biol. 11 (3), 1393-1401 (1991)			
COMMENT	9114187			
FEATURES	1705004			
source	Original			
	source text: Rat, cDNA to mRNA.			
	Location/Qualifiers			
	1..1500			
	/organism="Rattus norvegicus"			

```

/mol_type="mRNA"
/strain="fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AA41380.1"
/db_xref="GI:204733"
/translacion="MPEELIVSWPFIILSFQYRVVAGAPQPMHCACTAERLELCE
FPNSAFISRPACGCGCCPTCALPLGAAGVATRCAGLSGRALPGSPRLHALTGG
QGAVLBAAPATSLSGSHEEAKAAVASDELAESEMEBOLDPSFHMAASREI
OPTLMALSTSSMRARBITTLTKMKRPPCQRELKYVERLAQAQKADELYRYKLPR
CNKGFFHSKCETSLDGEAGLCWCVTPWSCKLPGLSDTRDENCHOYEVNQ"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match      18.1%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-10;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      341 CACAAACCAGCGAGCATTTGAACAACCTGCACACGGGCATTGTGCCAGAGAGCTGTGACAC 400
        |||||
DB       8 CACAAACCAGCGAGCATTTGAACAACCTGCACACGGGCATTGTGCCAGAGAGCTGTGACAC 67
QY      401 CACTTCGGCTACTAGCTA 418
        |||||
DB      68 CACTTCGGCTACTAGCTA 85

RESULT 10
AXI63782          5001 bp   DNA           linear    PAT 22-JUN-2001
LOCUS             AXI63782
DEFINITION        Sequence 46 from Patent WO0138579.
ACCESSION         AXI63782
VERSION           AXI63782.1 GI:14544878
KEYWORDS
SOURCE            Rattus norvegicus (Norway rat)
ORGANISM          Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE
AUTHORS           Gould-Rothberg,B.E., Dipippo,V.A., Ramsehl,T.M. and Gerwein,R.W.
TITLE             Method of identifying toxic agents using nsaid-induced differential
                gene expression in liver
JOURNAL           Patent: WO 0138579-A 46 31-MAY-2001;
                Curagen Corporation (US)
FEATURES
source            1..5001
                /organism="Rattus norvegicus"
                /mol_type="unassigned DNA"
                /db_xref="taxon:10116"

ORIGIN
Query Match      18.1%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 1.5e-10;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      341 CACAAACCAGCGAGCATTTGAACAACCTGCACACGGGCATTGTGCCAGAGAGCTGTGACAC 400
        |||||
DB       1 CACAAACCAGCGAGCATTTGAACAACCTGCACACGGGCATTGTGCCAGAGAGCTGTGACAC 60
QY      401 CACTTCGGCTACTAGCTA 418
        |||||
DB      61 CACTTCGGCTACTAGCTA 78

RESULT 11
RATIGPBA
```

341 CACAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 400

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAX Plate: 184 Row: e Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gt: 6981079.

```
location/Qualifiers
1. 1510
/organism="Rattus norvegicus"
/mol_type="mrna"
/db_xref="taxon:10116"
```

gene
/clone="WGC:93595 IMAGE:7129185"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_id="NIN MG_C236"
/lab_host="DH10B"
/note="Vector: pExpress1"
1. .1510
/gene="Igfbp1"
/note="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="locusid:25685"
/db_xref="ratmap:44422"
/db_xref="RGD:2872"
142. .960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH78889.1"
/db_xref="GI:50927647"
/db_xref="locusid:25685"
/db_xref="ratmap:44422"
/db_xref="RGD:2872"
/translation="MPEFLTVVSWPFLILSPQVAVVAGAPQPMHCAPCTAERLEICP
PYAPCPERSRAGCCCTCTCALPIGAAGVATACAGLSGRLPSPRPRLHATRG
OACVLEPAPATSSUSGQHEAPAAVAEDELAESEPMTEQLDSFILMAPEFED
OPLMNAISTYSMPAREITDLKKWKEPCQRELYVLERLAAQKAGDELYKFLPN
CKMGFYHSKQCEISLDGEAGLCWCVPWSGKKIPGSIETRDPMCHQYFVNVN"

ORIGIN

Query Match 15.5%; Score 65.4; DB 10; Length 1510;
Best Local Similarity 98.5%; Pred. No. 1.8e-07;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 352 CGAGATTGACATGCACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGGCTA 411
Db 1 CGAGATTGACATGCACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGGCTA 60

Qy 412 CTAGCTA 418
Db 61 CTATCTA 67

RESULT 13
AY095345 3886 bp DNA linear PRI 30-DEC-2002
LOCUS
DEFINITION Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
ACCESSION AY095345
VERSION AY095345.1 GI:20853764
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
REFERENCE
AUTHORS Kim,J.-J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
TITLE Regulation of insulin-like growth factor binding protein-1 promoter
activity by FKHR and HOTA10 in primate endometrial cells
JOURNAL Biol. Reprod. 68 (1), 24-30 (2003)
PUBMED 12453691
REFERENCE 2 (bases 1 to 3886)
AUTHORS Kim,J.-J., Jaffe,R.C. and Fazleabas,A.T.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

FEATURES
source
1. .3886
location/Qualifiers
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
<3655. .>3886
/gene="IGFBP-1"

mRNA
<3655. .>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655. .>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAH73273.1"
/db_xref="GI:20853765"
/translation="MSEVPVAVVVLITLVQVGVATASAPWQCAPSAAETALCPV
PASCSEVTRSGCCCPMCALPIGAAGVATAR"

ORIGIN

Query Match 14.8%; Score 62.6; DB 9; Length 3886;
Best Local Similarity 64.0%; Pred. No. 1.1e-06;
Matches 135; Conservative 0; Mismatches 59; Indels 17; Gaps 2;

CDS
Qy 206 CACAAGCAAAACAACTATTATTTGACACGGGATCTAGACGCTGCGTGAACATCAT 265
Db 3366 CACTAGCAAAACAACTATTATTTGACACGCTGCGTGAACATCAT 3425

Qy 266 TAAACC-----GTGCTGCCGAGCAGCCCTTATAAGCCCTGGTATGCCAGC 315
Db 3426 TAACTCTGTGCAAGTGGCGCGCTGTGCTTATATAGGCGCGCTGTGCCAGC 3485

Qy 316 CAGCATGTCCTACGCGCCGCGAGACAAACCCAGAGATGAAACATGCACACGCGC 375
Db 3486 AAGCATGCGCCACCGCCATCTCCATTCAGCAAGC-----ATTCGCCGCGCGCGCGC 3538

Qy 376 CATCTGCCAGAGAGCTGTGACCACTTC 406
Db 3539 CACCTCCAGAGAGACTGCGCCACCGCTCC 3569

RESULT 14
G67139 448 bp DNA linear STS 18-SEP-2000
LOCUS
DEFINITION IGFBP1 X1.1 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G67139
VERSION G67139.1 GI:10186730
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Cox,D.G., Bolliot,C. and Canzian,F.
TITLE Genome Survey
JOURNAL Unpublished (2000)
COMMENT
Contact: Federico Canzian
Genome Analysis Group
International Agency for Research on Cancer
150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France
Tel: +33-4-72738698
Fax: +33-4-72738388
Email: canzian@arc.fr
Primer A: TGCCTAGCAAAACAACTT
Primer B: GGGTAGCTTCGAGCAGC
STS size: 448
Protocol:
Template: 50 ng
Primer: 4 um each
dNTPs: 2 mM each
MgCl2: 1.5-2.5 mM
Tag: 0.05 units
Total Vol: 25 ul
Buffer: 1.5-2.5 mM
MgCl2: 50 mM
KCl: 20 mM
Tris-HCl: 8.4.
pH:

FEATURES
Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind
primer_bind
complement(431..448)
ORIGIN

Query Match 14.4%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 3.4e-06;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTTATTTTGAACAGGGGATCTTAGCAGCTGCGCTGACATCAT 265
DB 3 CACTAGCAAAACAACTTATTTTGAACAGCTCTAGCGCGCGCTGCAATCAT 62
QY 266 TAACCC-----GTGCTGCCGAGCGCCCTTATAGGCGCTGGGTATGGCCAGC 315
DB 63 TAACTCTCTGTCGCAAGTGCGCGCGCTGTGCTTATAGGTGCGCTGTGTCAGC 122
QY 316 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGCAGCATTTGAACACTGC -ACACGG 374
DB 123 GAGCATCGGCAACCGCATTC-----CATCAGCGAGCATCTGCCCGCGCGCGCG 174
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406
DB 175 CCACCTCCAGAGAGCACTGCGCACCGCTCC 206

RESULT 15

LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2394 from Patent WO229103.
ACCESSION AX409747
VERSION AX409747.1 GI:2142452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1..6128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN

Query Match 14.4%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.4e-06;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTTATTTTGAACAGGGGATCTTAGCAGCTGCGCTGACATCAT 265
DB 473 CACTAGCAAAACAACTTATTTTGAACAGCTCTAGCGCGCGCTGCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCGCCCTTATAGGCGCTGGGTATGGCCAGC 315
DB 533 TAACTCTCTGTCGCAAGTGCGCGCGCTGTGCTTATAGGTGCGCTGTGTCAGC 592
QY 316 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGCAGCATTTGAACACTGC -ACACGG 374
DB 593 GAGCATCGGCAACCGCATTC-----CATCAGCGAGCATCTGCCCGCGCGCGCG 644
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406

DB 645 CCACCTCCAGAGAGCACTGCGCACCGCTCC 676

Search completed: September 1, 2005, 03:27:00
Job time: 2637.79 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 3215.92 Seconds
(without alignments)
5006.706 Million cell updates/sec

Title: US-09-972-916B-6

Perfect score: 423

Sequence: 1 catggcgagcgagcgagcgaccc.....ttccgctactagctagcgacg 423

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	23.2	1013	2	BF236974
2	87.8	20.8	310	6	BY783538
3	87.8	20.8	310	6	BY783538
4	87.8	20.8	331	6	BY794229
5	86.8	20.5	354	6	BY771317
6	86.8	20.5	401	1	AI1785818
7	86.8	20.5	480	1	AI196314
8	86.8	20.5	706	1	AI1530146
9	86.8	20.5	765	1	AI1098594
10	86.8	20.5	799	1	AI1529939
11	86.8	20.5	811	6	CA478518
12	86.8	20.5	848	1	AI1790802
13	86.8	20.5	852	1	AI1528304
14	85.2	20.1	605	1	AI196154
15	82.2	19.4	360	1	AI1785039
16	81.4	19.2	846	7	CO573026
17	80.4	19.0	692	7	CV127049
18	79.4	18.8	713	7	CO560662
19	75.4	17.8	748	7	CO575629
20	73.8	17.4	694	7	CV117001
21	72.4	17.3	332	7	W30013
22	72.4	17.1	615	6	BB660958
23	72.4	17.1	618	6	CD561711
24	72.4	17.1	632	1	AI892189

25	67.4	15.9	488	1	AA674302
26	65.4	15.5	801	7	CK472246
27	64.4	15.2	840	7	CK473709
28	62.8	14.8	759	1	AA105355
29	61.6	14.6	269	2	BB604790
30	61.6	14.6	587	2	AA916227
31	60.8	14.4	1589	3	CR621807
32	60.8	14.4	1601	3	CR595377
33	59	13.9	545	1	AA060360
34	55.8	13.2	597	8	AZ652514
35	55.2	13.0	1167	3	CR667848
36	53.6	12.7	1128	3	CR657703
37	53.6	12.7	1128	3	CR671810
38	53.6	12.7	1159	3	CR663312
39	50.4	11.9	289	2	BE066031
40	48.8	11.5	234	2	AA086850
41	48.8	11.5	325	2	AA085818
42	47.2	11.2	307	2	AA086846
43	47.2	11.2	659	8	AZ840793
44	46.2	10.9	212	2	BE149547
45	45.8	10.8	270	2	AA075361

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4161889 5',
RNA sequence.
ACCESSION BF236974.1 GI:11150891
VERSION BF236974.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contract: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9443 row: j column: 02
High quality sequence stop: 581.
location/Qualifiers
1. 1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161889"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_L19"
/note="Organ: Liver; Vector: pCMV-SPORT6; Site 1: NciI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 23.2%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 3.7e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 308 TGGCAGCAGCATGTCTCCTCCCGCCGAGACACAAACGAGCATTTGAACACTG 367

Db	9	TGGGACAGCAGATGTGCCTACTGTCGCCCGGGAGAACAACACCAGCGAGATTGAACACTG	68
Oy	368	CACACGGCCATCTGCCAGAGAGCTGTGTGACCAACCACTTCCGCTACTAGCTA	418
LOCUS		BY783538	310 bp mRNA linear EST 23-MAR-2004
DEFINITION		BY783538 RIKEN full-length enriched, 17.5 days embryo whole body	
VERSION		Mus musculus cDNA clone L930176D05 5', mRNA sequence.	
KEYWORDS		BY783538	
SOURCE		BY783538.1 GI:39710177	
ORGANISM		Mus musculus (house mouse)	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		Carrollin,P., Waki,K., Shiraki,T., Konno,H., Shibata,K., Itoh,M., Aizawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S., Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Satou,K., Watabiki,A., Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A., Kiuchi,N., Yoshiki,A., Kusabe,M., Gustinich,S., Beisel,K., Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T., Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagioliini,M., Henschel,T.K., Brinkmeier,M., Camper,S., Hirota,J., Lombaerts,P., Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.	
TITLE		Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia	
JOURNAL MEDLINE		Genome Res. 13 (6B), 1273-1289 (2003)	
PUBMED		22703353	
COMMENT		12819125 Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp URL:http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.	
FEATURES		Location/Qualifiers	
source		1..310 organism="Mus musculus" mol_type="mRNA" strain="CS7BL/6J" db_xref="taxon:10090" clone="L930176D05" tissue.type="whole body" dev_stage="17.5 days embryo" clone_idb="RIKEN full-length enriched, 17.5 days embryo whole body"	
ORIGIN			
Query Match		20.8%; Score 87.8; DB 6; Length 310;	
Best Local Similarity		92.9%; Pred. No.2,8-e-14;	
Matches		92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
Oy	320	ATGTGTCACTGCCCGCGAGACAAACCCAGCGAGATTGAACACTGCACAGCGCCATC	379
Db	2	ATGTCCTACTCCCGCGAGACACACCCAGCGAGATTGAACACTGCACAGCGCCATC	61
Oy	380	TGCCAGAGAGCTGTGACCATCCTCCGCTACTAGCTA	418

[illegible]

RESULT 5	LOCUS	DEFINITION	ACCESSION
A1785818	A1785818	401 bp mRNA linear EST 02-JUL-1996	
	U1785818	Sugano mouse liver Mus musculus cDNA clone	
	IMAGE:1088569.5	similar to gp:XB1575 M.musculus mRNA for	
		insulin-like growth factor binding (MOUSE) ;, mRNA sequence.	
	A1785818		

LOCUS	DEFINITION	FEATURES
RESULT 6 A1196314	480 bp mRNA	linear EST 14-OCT-1998
A1196314	u71207.v1 Sugano mouse liver mRNA	musculus cDNA clone
IMAGE:1867828.5	similar to gb:459313.fna1	INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1	PRECURSOR (HUMAN); gb:461519	M. musculus
mRNA for insulin-like growth factor binding	(MOUSE); mRNA	sequence.

ACCESSION A1196314 GI:3748920
VERSION A1196314.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 480)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowleg,M., Dietrich,N., Dubuque,T., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972152
Seq primer: custom primer used
High quality sequence stop: 375.
Location/Qualifiers
1..480
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887828"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TTGTGGCTTCTGTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGTGAGCTCGACACA."
ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 480;
Best Local Similarity 92.9%; Pred. No. 5,9e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 321 TGTGTCACCTGCCCCGAGACACAAACCGAGAGATTTGAACACTGCACAGGCACT 380
DB 1 TGTGTCACCTGCCCCGAGACACACACCCAGGAGATTGAACACTGCACAGGCGCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
RESULT 7
A1530146 706 bp mRNA linear EST 18-MAR-1999
LOCUS A1530146
DEFINITION u189f09.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1889609 5' similar to gb:M5316 tml INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
ACCESSION A1530146 GI:4444281
VERSION A1530146.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 706)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:973933
Seq primer: custom primer used
High quality sequence stop: 479.
Location/Qualifiers
1..706
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889609"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TTGTGGCTTCTGTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGTGAGCTCGACACA."
ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 706;
Best Local Similarity 92.9%; Pred. No. 6,4e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 321 TGTGTCACCTGCCCCGAGACACAAACCGAGAGATTTGAACACTGCACAGGCACT 380
DB 1 TGTGTCACCTGCCCCGAGACACACACCCAGGAGATTGAACACTGCACAGGCGCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
RESULT 8
A1098594 765 bp mRNA linear EST 20-AUG-1998
LOCUS A1098594
DEFINITION u61e07.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1481988 5' similar to gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION A1098594

VERSION A1098594.1 GI:3448119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watscn.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 395.
Location/Qualifiers
1..765
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1481988"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMB18-F13; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG]; digested and cloned into distinct DraIII sites of the pMB18-F13 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCTGCTTAAGAAGCTGG and 3' end primer CGACTGCAGCTCGAGCA."

ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 765;
Best Local Similarity 92.9%; Pred. No. 6.5e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTTCACCTGCCCGCGAGACACAAACCCAGGACATTGAACACTGCACAGGCGCATCT 380
DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGGACATTGAACACTGCACAGGCGCATCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCATTCCTCGCTACTACTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCATTCCTCGCTACTACTA 98

RESULT 9
A1530313 785 bp mRNA linear EST 18-MAR-1999
LOCUS A1530313
DEFINITION u187c09.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
VERSION A1530313.1 GI:4444448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watscn.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974109
Seq primer: custom primer used
High quality sequence stop: 459.
Location/Qualifiers
1..785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMB18-F13; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG]; digested and cloned into distinct DraIII sites of the pMB18-F13 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCTGCTTAAGAAGCTGG and 3' end primer CGACTGCAGCTCGAGCA."

ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 785;
Best Local Similarity 92.9%; Pred. No. 6.6e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTTCACCTGCCCGCGAGACACAAACCCAGGACATTGAACACTGCACAGGCGCATCT 380
DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGGACATTGAACACTGCACAGGCGCATCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCATTCCTCGCTACTACTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCATTCCTCGCTACTACTA 98

RESULT 10
A1529939 799 bp mRNA linear EST 18-MAR-1999
LOCUS A1529939
DEFINITION u187c09.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1889392 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI529939 GI:4444074
 VERSION AI529939.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 799)
 REFERENCE 1 (bases 1 to 799)
 AUTHORS Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Maria M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MG1:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 Location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:189392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME185-Fu3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGACCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGTGCTACTGTG], digested and cloned into distinct DraIII sites of the pME185-Fu3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTGTCTCTTAAGCTGTG and 3' end primer CGACCTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 20.5%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 6.6e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 321 TGTGTCACTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGGCCATCT 380
 Db 1 TGTGTCACTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGGCCGCTCT 60
 Oy 381 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 418
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 98
 RESULT 11
 LOCUS CA478518 811 bp mRNA linear EST 09-MAR-2004
 DEFINITION AGENCOURT 10789306 NIH MGC 152 Mus musculus cDNA clone
 IMAGE:6766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 811)
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Plate: LLM00138 row: m column: 23
 High quality sequence stop: 536.
 Location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCCTGACAGCCGCGCACCACTTGAACAAGAGCTGGGTTTTTTTTTTT TTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen, library amplification 16 cycles, library constructed by Mark Bittinger in the Bradfield Laboratory (McArdle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 20.5%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 6.6e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 321 TGTGTCACTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGGCCATCT 380
 Db 39 TGTGTCACTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGGCCGCTCT 98
 Oy 381 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 418
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 136
 RESULT 12
 LOCUS A1790802 848 bp mRNA linear EST 02-JUL-1999
 DEFINITION uk28b10.y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 IMAGE:1970299 5' similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 848)
 REFERENCE 1 (bases 1 to 848)
 AUTHORS Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28b10.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039

FEATURES

source

High quality sequence stop: 514.
Location/Qualifiers

1..848
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse kidney mklia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGTCTCTAAAGCTGGG and 3' end primer CGACTGCGAGCTCGAGCACA."

ORIGIN

Query Match 20.5%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 6.7e-14;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGGTCCACGTCCCGCGGAGACACAAACCGAGGAGCATGAAACATGACACAGGCGCATCT 380

DB 1 TGGTCCACGTCCCGCGGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGCGCTCT 60

QY 381 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 418

DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 98

RESULT 13

AI528304 852 bp mRNA linear EST 18-MAR-1999

LOCUS AI528304 u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone

DEFINITION IMAGE:1890210 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI528304
VERSION AI528304.1 GI:4442439

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534

FEATURES

source

High quality sequence stop: 478.
Location/Qualifiers

1..852
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct draiii sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGTCTCTAAAGCTGGG and 3' end primer CGACTGCGAGCTCGAGCACA."

ORIGIN

Query Match 20.5%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 6.7e-14;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGGTCCACGTCCCGCGGAGACACAAACCGAGGAGCATTTGAACACTGCACAGGCGCATCT 380

DB 1 TGGTCCACGTCCCGCGGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGCGCTCT 60

QY 381 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 418

DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 98

RESULT 14

AI196154 605 bp mRNA linear EST 14-OCT-1998

LOCUS AI196154 u195d08.y1 Sugano mouse liver mla Mus musculus cDNA clone

DEFINITION IMAGE:1887663 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI196154
VERSION AI196154.1 GI:3748760

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMNI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:971987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

```
1..605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/notes="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
primer CGACTGCGACTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 20.1%; Score 85.2; DB 1; Length 605;
Pred. No. 1.8e-13;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Oy 321 TGGTCCACTGCGCGGAGACACAAACCGAGAGATTGAACACTGACACGCGCATCT 380
|||||
Db 1 TGGTCCACTGCGCGGAGACACACACCCAGGACATTGAACACTGACACGCGTCTCT 60
|||||
Oy 381 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTTA 418
|||||
Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCACTACTACTTA 98
|||||
```

RESULT 15

A1785039

u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone

LOCUS

IMAGE:1888018 5' similar to gb:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE)"; mRNA sequence.

ACCESSION

A1785039

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

380 bp mRNA linear EST 02-JUL-1999
A1785039
u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888018 5' similar to gb:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE)"; mRNA sequence.
A1785039
A1785039.1 GI:5332755
EST.
Mus musculus
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 380)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

```
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/notes="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
primer CGACTGCGACTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 19.4%; Score 82.2; DB 1; Length 380;
Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Oy 324 TCCACTGCGCGCGGAGACACAAACCGAGAGATTGAACACTGACACGCGCATCTGCC 383
|||||
Db 1 TCCACTGCGCGCGGAGACACACACCCAGGACATTGAACACTGCATACGCGCCGTCTGCC 60
|||||
Oy 384 CAGAGAGCTGTGACCAACCACTTCGGCTACTACTTA 418
|||||
Db 61 CAGAGAGCTGTGACCAACCACTTCGCACTACTACTTA 95
|||||
```

Search completed: September 1, 2005, 06:56:53
Job time : 3217.09 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 ; Search time 1231.96 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-6

Perfect score: 423

Sequence: 1 catggggcgcacggggcactc.....ttccgctactagctagccgc 423

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	423	100.0	423	9	US-09-972-916A-6
2	270	63.8	270	9	US-09-972-916A-3
3	223.6	52.9	372	9	US-09-972-916A-5
4	219.4	51.9	321	9	US-09-972-916A-4
5	219	51.8	219	9	US-09-972-916A-2
6	154	36.4	372	9	US-09-972-916A-5
7	103	24.3	321	9	US-09-972-916A-4

8	76.4	18.1	1500	9	US-09-917-800A-1608	Sequence 1608, Ap
9	76.4	18.1	1500	17	US-10-388-934-5	Sequence 5, Appli
10	76.4	18.1	1500	17	US-10-191-803-73	Sequence 73, Appli
11	76.4	18.1	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
12	60.8	14.4	6128	9	US-09-880-107-2393	Sequence 2393, Ap
13	60.8	14.4	6128	22	US-10-756-149-1484	Sequence 1484, Ap
14	60.8	14.4	9173	22	US-10-893-315-126	Sequence 126, App
15	60.8	14.4	9174	22	US-10-893-315-160	Sequence 160, App
16	55.2	13.0	13011	17	US-10-388-934-36	Sequence 36, Appli
17	55.2	13.0	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
18	51	12.1	51	9	US-09-972-916A-1	Sequence 1, Appli
19	48.4	11.4	3191	16	US-10-029-386-0306	Sequence 20306, A
20	46.2	10.9	12850	14	US-10-017-122-1	Sequence 17872, A
21	45.6	10.8	586	16	US-10-029-386-5186	Sequence 5186, Ap
22	43.6	10.3	90556	21	US-10-741-600-17972	Sequence 17972, A
23	42.8	10.1	423	9	US-09-972-916A-6	Sequence 6, Appli
24	41	9.7	3037	20	US-10-363-345A-40659	Sequence 40659, A
25	41	9.7	3037	20	US-10-363-345A-40660	Sequence 40660, A
26	41	9.7	3037	21	US-10-363-483A-40659	Sequence 40659, A
27	41	9.7	3037	21	US-10-363-483A-40660	Sequence 40660, A
28	40.8	9.6	28000	16	US-10-091-625-11	Sequence 11, Appli
29	40.8	9.6	28000	17	US-10-096-399A-11	Sequence 11, Appli
30	40.8	9.6	28000	17	US-10-461-668-11	Sequence 11, Appli
31	40.8	9.6	28000	17	US-10-363-345A-40659	Sequence 381, App
32	40.8	9.6	148083	22	US-10-756-149-4571	Sequence 4571, Ap
33	40.2	9.5	31703	17	US-10-085-117-172	Sequence 17, App
34	39.4	9.2	201	21	US-10-741-600-69952	Sequence 69952, A
35	39	9.3	201	20	US-10-719-993-35505	Sequence 35505, A
36	38.6	9.1	34217	20	US-10-719-993-6865	Sequence 6865, Ap
37	38.4	9.1	11009	9	US-09-845-583-1	Sequence 1, Appli
38	38.4	9.1	11009	14	US-10-037-182-3	Sequence 3, Appli
39	38.4	9.1	11009	21	US-10-764-420-947	Sequence 947, App
40	37.8	8.9	201	20	US-10-719-993-35504	Sequence 35504, A
41	37.8	8.9	201	20	US-10-719-993-35509	Sequence 35509, A
42	37.8	8.9	905	20	US-10-363-345A-40663	Sequence 40663, A
43	37.8	8.9	905	20	US-10-363-345A-40664	Sequence 40664, A
44	37.8	8.9	905	21	US-10-363-483A-40664	Sequence 40664, A
45	37.8	8.9	905	21	US-10-363-483A-40664	Sequence 40664, A

ALIGNMENTS

RESULT 1
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match 100.0%; Score 423; DB 9; Length 423;
Bee Local Similarity 100.0%; Pred. No. 2, 2e-116; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGGCGCAGCGGGGCACTCCCGTGTTCTTGACTTGGCCCCAGTGTACATGAGGCGC 60
DB 1 CATGGGCGCAGCGGGGCACTCCCGTGTTCTTGACTTGGCCCCAGTGTACATGAGGCGC 60
QY 61 ACGGGGCACTCCCGTGTTCTTGACTTGGCCCCAGTGTACATGAGGCGCAGGCGC 120

```
Db 61 AGGGGCACTCCCGTGTCTGACTGCGCCCGCAGGTACATGGGGCAGCGGGCAC 120
Qy 121 TCCCGGTGTTCTGACTCTGCGCCCGCAGGTATCATGGGGCAGCGGGCACTCCGTGT 180
Db 121 TCCCGGTGTTCTGACTCTGCGCCCGCAGGTATCATGGGGCAGCGGGCACTCCGTGT 180
Qy 181 TCCGTGACTCTGGCCCCCAGGTATCACAAGCAAAACAACTTATTTTGAACAGGGGAT 240
Db 181 TCCGTGACTCTGGCCCCCAGGTATCACAAGCAAAACAACTTATTTTGAACAGGGGAT 240
Qy 241 CCTAGACGCTGCTCCCTGACATCATTAACCCGTGCTGCCAGCCAGCCCTTCATAAGCC 300
Db 241 CCTAGACGCTGCTCCCTGACATCATTAACCCGTGCTGCCAGCCAGCCCTTCATAAGCC 300
Qy 301 CTGGGTATGGCAGCAGCATGTGTCACCTGCGCCGCGAGACACAAACCCAGGAGCATTTG 360
Db 301 CTGGGTATGGCAGCAGCATGTGTCACCTGCGCCGCGAGACACAAACCCAGGAGCATTTG 360
Qy 361 AACACTGCACACGGGCATCTGCGCCAGAGAGCTGTGACCACTTCCGCTACTAGTAGC 420
Db 361 AACACTGCACACGGGCATCTGCGCCAGAGAGCTGTGACCACTTCCGCTACTAGTAGC 420
Qy 421 CGC 423
Db 421 CGC 423
```

RESULT 2
US-09-972-916A-3

```
/ Sequence 3, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 3
/ LENGTH: 270
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-3
```

Query Match 63.8%; Score 270; DB 9; Length 270;

Best Local Similarity 100.0%; Pred. No. 2.7e-83;

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 154 CATGGGCGCAGCGGGCACTCCCGTGTCTGACTCTGGCCCCCAGGTATCACAAGCA 213
Db 1 CATGGGCGCAGCGGGCACTCCCGTGTCTGACTCTGGCCCCCAGGTATCACAAGCA 60
Qy 214 AAACAACTTATTTTGAACAGGGGATCTAGCAGCTGCCCTGCATCATTAACCCGT 273
Db 61 AAACAACTTATTTTGAACAGGGGATCTAGCAGCTGCCCTGCATCATTAACCCGT 120
Qy 274 GCTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCATGCGCC 333
Db 121 GCTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCATGCGCC 180
Qy 334 GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 393
Db 181 GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 240
Qy 394 TGACCAACACTTCCGCTACTAGTAGCGGC 423
Db 241 TGACCAACACTTCCGCTACTAGTAGCGGC 270
```

RESULT 3
US-09-972-916A-5

```
/ Sequence 5, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 5
/ LENGTH: 372
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-5
```

Query Match 52.9%; Score 223.6; DB 9; Length 372;

Best Local Similarity 76.2%; Pred. No. 3.9e-67;

Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

```
Qy 34 ACTGTGCCCCCAGGTACATGGGGCAGCAGGGGCACTCCGTGTCTGACTCTGAGCC 93
Db 2 AACTGGGGGCCAGAGTCCAGATCCACAGGAGATGCCCTCGTCCCATGTACACTGGGG 61
Qy 94 CCAGGTACATGGGGCAGCAGGGGCACTCCGTGTCTGACTCTGAGCCCGCAGTETA 153
Db 62 GCCAGAGTCCAGAGAACACGGGAGTGCCCTGTGCGCCATGTACATGGGGCCAGAGTC 121
Qy 154 CATGGGCGCAGCGGCACTCCCGTGTCTGACTCTGAGCTGGCCCCCAGGTATCACAAGCA 213
Db 122 CAGAACCAACGAGGTGCCCGTGTGCGCCATG-----TCACAAGCA 162
Qy 214 AAACAACTTATTTTGAACAGGGGATCTAGCAGCTGCCCTGCATCATTAACCCGT 273
Db 163 AAACAACTTATTTTGAACAGGGGATCTAGCAGCTGCCCTGCATCATTAACCCGT 222
Qy 274 GCTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCATGCGCC 333
Db 223 GCTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGTGTCATGCGCC 282
Qy 334 GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 393
Db 283 GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 342
Qy 394 TGACCAACACTTCCGCTACTAGTAGCGGC 423
Db 343 TGACCAACACTTCCGCTACTAGTAGCGGC 372
```

RESULT 4
US-09-972-916A-4

```
/ Sequence 4, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 4
/ LENGTH: 321
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
```

OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 51.8%; Score 219.4; DB 9; Length 321;
Best Local Similarity 99.5%; Pred. No. 1.1e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 TATCAAGCAAAACAACTATTATTGAACAGCGGGATCTTACAGAGCTGCGCTGACAT 262
DB 101 TGTCAAGCAAAACAACTATTATTGAACAGCGGGATCTTACAGAGCTGCGCTGACAT 160
QY 263 CATTAACCCGTGCTGCGGAGCCAGCCCTTCAATAGGCGCTGATAGCCAGCCAGCATG 322
DB 161 CATTAACCCGTGCTGCGGAGCCAGCCCTTCAATAGGCGCTGATAGCCAGCCAGCATG 220
QY 323 GTTCACTGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGC 382
DB 221 GTTCACTGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGC 280
QY 383 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 423
DB 281 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 321

RESULT 5

US-09-972-916A-2
Sequence 2, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 2

LENGTH: 219

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

US-09-972-916A-2

Query Match 51.8%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e-65;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 TCACAGCAAAACAACTATTATTGAACAGCGGGATCTTACAGAGCTGCGCTGACATCA 264
DB 1 TCACAGCAAAACAACTATTATTGAACAGCGGGATCTTACAGAGCTGCGCTGACATCA 60
QY 265 TTAACCCGTGCTGCGGAGCCAGCCCTTCAATAGGCGCTGATAGCCAGCCAGCATGT 324
DB 61 TTAACCCGTGCTGCGGAGCCAGCCCTTCAATAGGCGCTGATAGCCAGCCAGCATGT 120
QY 325 CCATGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCC 384
DB 121 CCATGCGCGCGGAGACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCC 180
QY 385 AGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 423
DB 181 AGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 219

RESULT 6

US-09-972-916A-5/c
Sequence 5, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

CURRENT FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 5

LENGTH: 372

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 36.4%; Score 154; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 5.9e-43;

Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ACATGGGCGCAAGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGCG 110
DB 154 ACATGGGCGCAAGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGCG 95
QY 111 CACGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 170
DB 94 CACGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 35
QY 171 CTCCTGTGTTCTTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 204
DB 34 CTCCTGTGTTCTTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 1

RESULT 7

US-09-972-916A-4/c
Sequence 4, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 4

LENGTH: 321

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 24.3%; Score 103; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.9e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ACATGGGCGCAAGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGCG 110
DB 103 ACATGGGCGCAAGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGCG 44
QY 111 CACGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 153
DB 43 CACGGGGCACTCCCGTGTCTGACTGTGGCCCCCAGTGTACATGGGGCGGCA 1

RESULT 8

US-09-917-800A-1608
Sequence 1608, Application US/09917800A
Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Casale, Arthur

Query Match	18.1%;	Score 76.4;	DB 17;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 7.9e-16;		

RESULT 1
US-10-152-319A-1613
Sequence 1613, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335

;; PRIOR FILING DATE: 2001-05-22
;; PRIOR APPLICATION NUMBER: US 60/297,523
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/298,925
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: US 60/303,810
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/303,807
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/303,808
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/315,047
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: US 60/324,928
;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: US 60/330,867
;; PRIOR FILING DATE: 2001-11-01
;; PRIOR APPLICATION NUMBER: US 60/330,462
;; PRIOR FILING DATE: 2001-10-22
;; Remaining prior Application data removed - See File Wrapper or PAM.
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 1613
;; LENGTH: 1500
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; FEATURE:
;; OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match 18.1%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7.9e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATTTGCCAGAGAGCTGTACAC 400
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATTTGCCAGAGAGCTGTACAC 67

QY 401 CACTTCGCTACTAGCTA 418
DB 68 CACTTCGCTACTACTTA 85

RESULT 12
US-09-880-107-2393
;; Sequence 2393, Application US/09880107
;; Patent No. US20020142981A1
;; GENERAL INFORMATION:
;; APPLICANT: Horne, Darci T.
;; APPLICANT: Vockley, Joseph G.
;; APPLICANT: Scherf, Uwe
;; APPLICANT: Gene Logic, Inc.
;; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;; FILE REFERENCE: 44921-5028-WO
;; CURRENT APPLICATION NUMBER: US/09/880,107
;; PRIOR FILING DATE: 2001-06-14
;; PRIOR APPLICATION NUMBER: US 60/211,379
;; PRIOR FILING DATE: 2000-06-14
;; PRIOR APPLICATION NUMBER: US 60/237,054
;; PRIOR FILING DATE: 2000-10-02
;; NUMBER OF SEQ ID NOS: 3950
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 2393
;; LENGTH: 6128
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
US-09-880-107-2393

Query Match 14.4%; Score 60.8; DB 9; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAAGCAAAACAATTATTTGACACGGGGATCTTAGACGCTGCCCTGACAAATCAT 265
DB 473 CACTGCAAAACAATTATTTGACACTAGCTCTTAGCGTGGCGCTGCCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCAGCCCTTCAATAGGCCCTGGATAGCCAGC 315
DB 533 TAACCTCGTGCAAGTGGCGCGCGCTGTGCTTTAATAGTGGCGCTGTGTCAGC 592
QY 316 CAGCATGTCTCACTGCGCGCGCGAGACAAACCCAGCAGATTGAACATGCG-ACACGG 374
DB 533 GAGCATGCGCCACCCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCG 644
QY 375 CCATCTGCCAGAGAGCTGTGACCAACCACTTC 406
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 13
US-10-756-149-1484
;; Sequence 1484, Application US/10756149
;; Publication No. US20050181375A1
;; GENERAL INFORMATION:
;; APPLICANT: Aziz, Natsasha
;; APPLICANT: Zlotnick, Albert
;; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;; FILE REFERENCE: file
;; CURRENT APPLICATION NUMBER: US/10/756,149
;; PRIOR FILING DATE: 2004-01-12
;; NUMBER OF SEQ ID NOS: 5818
;; SOFTWARE: Patentln version 3.2
;; SEQ ID NO 1484
;; LENGTH: 6128
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-756-149-1484

Query Match 14.4%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAAGCAAAACAATTATTTGACACGGGGATCTTAGACGCTGCCCTGACAAATCAT 265
DB 473 CACTGCAAAACAATTATTTGACACTAGCTCTTAGCGTGGCGCTGCCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCAGCCCTTCAATAGGCCCTGGATAGCCAGC 315
DB 533 TAACCTCGTGCAAGTGGCGCGCGCTGTGCTTTAATAGTGGCGCTGTGTCAGC 592
QY 316 CAGCATGTCTCACTGCGCGCGCGAGACAAACCCAGCAGATTGAACATGCG-ACACGG 374
DB 593 GAGCATGCGCCACCCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCG 644
QY 375 CCATCTGCCAGAGAGCTGTGACCAACCACTTC 406
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 14
US-10-893-315-126
;; Sequence 126, Application US/10893315
;; Publication No. US20050147987A1
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
;; FILE REFERENCE: CLO00786
;; CURRENT APPLICATION NUMBER: US/10/893,315
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: 60/231,397
;; PRIOR FILING DATE: 2000-09-08

```
/ NUMBER OF SEQ ID NOS: 2172
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 126
/ LENGTH: 9173
/ TYPE: DNA
/ ORGANISM: Human
US-10-893-315-126
```

```
Query Match          14.4%; Score 60.8; DB 22; Length 9173;
Best Local Similarity 66.5%; Pred. No. 3.5e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
```

```
QY 206 CACAAGCAAAACAACTTATTTTGAACAGGGGATCTTAGACGCTGCGCTGACATCAT 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1877 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGCGGCGCTGCCATCAT 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 TAAACC-----GTGCTGCCGAGCCGCGCTTTCATAGGCGCTTGATGCGCAGC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1937 TAACTCTCTGTGCAAGTGCGCGGCGCTGTGCTTTTAAGTGTGCGGCTGTGCCAGC 1996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 CAGCATGTGTCACGTGCGCGCGAGACACAACCCGAGCATTTGAACACTGC-ACAGCG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1997 GAGCATGCGCCACCGCCATCC-----CATCCAGCGAGCATGTGCCGCGCGCGCGC 2048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2049 CCACCTCTCCAGAGAGCACTGGCCACCGCTCC 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

```
US-10-893-315-160
/ Sequence 160, Application US/10893315
/ Publication No. US20050147987A1
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: C1000786
/ CURRENT APPLICATION NUMBER: US/10/893,315
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: 60/231,397
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 2172
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 160
/ LENGTH: 9174
/ TYPE: DNA
/ ORGANISM: Human
US-10-893-315-160
```

```
Query Match          14.4%; Score 60.8; DB 22; Length 9174;
Best Local Similarity 66.5%; Pred. No. 3.5e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
```

```
QY 206 CACAAGCAAAACAACTTATTTTGAACAGGGGATCTTAGACGCTGCGCTGACATCAT 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1877 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGCGGCGCTGCCATCAT 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 TAAACC-----GTGCTGCCGAGCCGCGCTTTCATAGGCGCTTGATGCGCAGC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1937 TAACTCTCTGTGCAAGTGCGCGGCGCTGTGCTTTTAAGTGTGCGGCTGTGCCAGC 1996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 CAGCATGTGTCACGTGCGCGCGAGACACAACCCGAGCATTTGAACACTGC-ACAGCG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1997 GAGCATGCGCCACCGCCATCC-----CATCCAGCGAGCATGTGCCGCGCGCGCGCG 2048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2049 CCACCTCTCCAGAGAGCACTGGCCACCGCTCC 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 1, 2005, 16:44:32

Job time : 1233.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 133.592 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916b-6

Perfect score: 423
Sequence: 1 catggcgccagcgggcgactc.....ttccgctactagctagccgc 423

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C_COMB.seq: *
6: /cgn2_6/prodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.8	14.4	9173	4 US-09-949-001-30	Sequence 30, Appl
2	60.8	14.4	9174	4 US-09-949-001-36	Sequence 36, Appl
3	55.2	13.0	194	4 US-08-845-140-1	Sequence 1, Appl
4	55.2	13.0	13011	2 US-08-791-849A-14	Sequence 14, Appl
5	44.8	10.6	40586	4 US-09-949-016-16965	Sequence 16965, A
6	41	9.7	767677	4 US-09-949-016-12147	Sequence 12147, A
7	41	9.7	767677	4 US-09-949-016-17361	Sequence 17361, A
8	40.2	9.5	15108	4 US-09-949-016-11786	Sequence 11786, A
9	40.2	9.5	15108	4 US-09-949-016-17205	Sequence 17205, A
10	38.4	9.1	11009	4 US-09-845-583A-1	Sequence 1, Appl
11	37.2	8.8	35471	4 US-09-949-016-12957	Sequence 12957, A
12	36.8	8.7	30245	4 US-09-949-016-12208	Sequence 12208, A
13	36.8	8.7	30245	4 US-09-949-016-13550	Sequence 13550, A
14	36.6	8.6	3054	3 US-09-484-970B-138	Sequence 138, App
15	36.4	8.6	134008	4 US-09-949-016-13841	Sequence 13841, A
16	36	8.5	16520	4 US-09-949-016-14710	Sequence 14710, A
17	36	8.5	16520	4 US-09-949-016-15394	Sequence 15394, A
18	35.4	8.4	1503	4 US-09-902-540-1970	Sequence 1970, Ap
19	35.4	8.4	16423	4 US-09-902-540-1120	Sequence 1120, Ap
20	35	8.3	601	4 US-09-949-016-19194	Sequence 19194, A
21	35	8.3	601	4 US-09-949-016-21963	Sequence 21963, A
22	35	8.3	2393	4 US-09-949-016-2153	Sequence 2153, Ap
23	35	8.3	2408	4 US-09-949-016-74	Sequence 74, Appl
24	35	8.3	7561	4 US-09-949-016-11816	Sequence 11816, A
25	35	8.3	7561	4 US-09-949-016-13895	Sequence 13895, A
26	34	8.0	749	3 US-09-257-583-12	Sequence 12, Appl
27	33.8	8.0	601	4 US-09-949-016-86816	Sequence 86816, A

C 28	33.8	8.0	601	4 US-09-949-016-86817	Sequence 86817, A
C 29	33.8	8.0	601	4 US-09-949-016-86818	Sequence 86818, A
C 30	33.8	8.0	601	4 US-09-949-016-86819	Sequence 86819, A
C 31	33.8	8.0	601	4 US-09-949-016-86820	Sequence 86820, A
C 32	33.8	8.0	601	4 US-09-949-016-86821	Sequence 86821, A
C 33	33.8	8.0	601	4 US-09-949-016-86822	Sequence 86822, A
C 34	33.8	8.0	247781	4 US-09-949-016-14193	Sequence 14193, A
C 35	33.6	7.9	601	4 US-09-949-016-21962	Sequence 21962, A
C 36	33.6	7.9	601	4 US-09-949-016-21963	Sequence 21963, A
C 37	33.6	7.9	601	4 US-09-949-016-84283	Sequence 84283, A
C 38	33.6	7.9	601	4 US-09-949-016-84284	Sequence 84284, A
C 39	33.6	7.9	14115	4 US-09-949-016-17490	Sequence 17490, A
C 40	33.6	7.9	455726	4 US-09-949-016-14157	Sequence 14157, A
C 41	33.6	7.9	481115	4 US-09-949-016-11940	Sequence 11940, A
C 42	33.4	7.9	601	4 US-09-949-016-35649	Sequence 35649, A
C 43	33.4	7.9	601	4 US-09-949-016-177807	Sequence 177807, A
C 44	33	7.8	57978	4 US-09-949-016-16667	Sequence 16667, A
C 45	32.8	7.8	1503	4 US-09-902-540-1970	Sequence 1970, Ap

ALIGNMENTS

```
RESULT 1
US-09-949-001-30
; Sequence 30, Application US/09949001
; Patent No. 6825336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-30

Query Match      14.4%; Score 60.8; DB 4; Length 9173;
Best Local Similarity 66.5%; Pred. No. 2e+09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      206 CACAGCAAAACAACTTATTTTGAACGCGGATCTAGCACGCTGCCCTGACATCAT 265
      1877 CACTGCAAAACAACTTATTTTGAACACTAGCTCTAGCGCTGGCGCTGCCATCAT 1936
      266 TAACCC-----GTGCTCCGAGCCAGCCCTTCAATGAGCCCTGGATGCGCAGC 315
      1937 TAACTCTCTGCGAGTAGTGCGCGGCTGTGCTTTATAGTGTCCGCTGTCTCAGC 1996
      316 CAGCATGTTCACCTGCGCGCGAGACACAAACCGAGCAGCATTTGAACACTGC-ACACG 374
      1997 GAGCATGCGCGCACCGCATTC-----CATTCAGCAGCATCTGCGCGCGCGCGC 2048
      375 CCATCTGCCGAGAGCTGTGACCCCACTTC 406
      2049 CCACCTCCAGAGAGCACTGCCACCGCTCC 2080

RESULT 2
US-09-949-001-36
; Sequence 36, Application US/09949001
; Patent No. 6825336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
```

FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match
Best Local Similarity 66.5%; Pred. No. 2e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

14.4%; Score 60.8; DB 4; Length 9174;
Best Local Similarity 66.5%; Pred. No. 2e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAAGCAAAACAACTTATTTTGAACACGCGGATCTAGACAGCTGCGCTTGACATCAT 265
DB 1877 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCAATCAT 1936
QY 266 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAATAGGCTTGCGGTATGCGCAGC 315
DB 1937 TAACCTCCTGTGCAAGTGGCGGCTGTGCTTATAGGTGCGCGCTGTGTCAGC 1996
QY 316 CAGCATGTGCACTGCGCGCGGAGACAAACCCAGCGAGCTTGAACACTGC-ACAGCG 374
DB 1997 GAGCATGCGCAGCCGCTATCC-----CATTCAGCGAGCATGTGCGCGCGCGCGCGC 2048
QY 375 CCATGTGCCAGAGAGCTGTGACCAACCATCTC 406
DB 2049 CCACCTCCCAAGAGCACTGGCGCACGCTCC 2080

RESULT 3
US-08-945-140-1
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816

TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match
Best Local Similarity 88.2%; Pred. No. 2.4e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

13.0%; Score 55.2; DB 3; Length 194;
Best Local Similarity 88.2%; Pred. No. 2.4e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CATGGGCGCAGGGGCACTCCCGTGTCTCTGACTCTGCGCCCGAGTACATGGGCGC 60
DB 11 CATGGGCGCAGGGGCACTCCCGTGTCTCTGACTCTGCGCCCGAGTACATGGGCGC 70
QY 61 ACGGGGCA 68
DB 71 CGTTGGCA 78

RESULT 4
US-08-791-849A-14
Sequence 14, Application US/08791849A
Patent No. 5914449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
TITLE OF INVENTION: Lipid Content in Plant Seed
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Penack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077)


```
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077)
FEATURE:
NAME/KEY: Intron
LOCATION: join(3219..3765, 3949..5916, 6009..6151,
6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297)
LOCATION: 9480..10162)
US-08-791-849A-14
```

```
Query Match 13.0%; Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.6e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CATGGCGCAGCGGAGCACTCCGTTGCTCTGACTCTGGCCCCCAGTGTACATGGGCGC 60
DB 3021 CATGGCGCAGCGGAGCACTCCGTTGCTCTGACTCTGGCCCCCAGTGTACATGGGCTTC 3080
QY 61 ACGGGGCA 68
DB 3081 CGTTGGCA 3088
```

```
RESULT 5
US-09-949-016-16965/C
```

```
; Sequence 16965; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16965
; LENGTH: 40586
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16965
```

```
Query Match 10.6%; Score 44.8; DB 4; Length 40586;
Best Local Similarity 51.5%; Pred. No. 0.00073;
Matches 103; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
QY 4 GGGCGCAGCGGAGCACTCCGTTGCTCTGACTCTGGCCCCCAGTGTACATGGGCGCAG 63
DB 26135 GGGGTCGCTGGGAGGAGGTGGTGCACGAGGTGTCCCAAGTGTCTCAGGGGCAAG 26076
QY 64 GGGCACTCCCGTTGCTCTGACTCTGGCCCCCAGTGTACATGGGCGCAGGGGCACTCC 123
DB 26075 GGGGAGGAGAGGTGGTGCAGGGGAGGTGTCCCAAGTGTCTCAGGGGCTGGGGGAGGGA 26016
QY 124 CGTGTCTCTGACTCTGGCCCCCAGTGTACATGGGCGCAGGGGCACTCCCGTTGCTTC 183
DB 26015 GGTGGTGTGCAAGTGAAGTGTCCCAAGTGTCTCGGGGGCTGGGGGAGGAGAGTGGGTGC 25956
QY 184 TGGACTCTGGCCCCCAGTGT 203
DB 25955 AGGCGAGGTGTCCCAAGT 25936
```

```
RESULT 6
US-09-949-016-12147
```

```
; Sequence 12147; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(76767)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147
```

```
Query Match 9.7%; Score 41; DB 4; Length 76767;
Best Local Similarity 51.4%; Pred. No. 0.05;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
```

```
QY 14 GGCATCCCCGTGTTCTCTGACTCTGGCCCCCAGTGTACATGGGCGCAGGGGCACTCCC 73
DB 562328 GGGAGTCCCGTGTCTCTTCTCTCAGGGCCCGTAGTACTGTAGGAGTGGAGTCCC 562387
QY 74 GGTGTTCTGAGCTCTGCCCCCAGTGTACATGGGCGCAGGGGCACTCCCGTTGCTCT 133
DB 562388 GTGATCTCAGACTAGAGGCTCCGTGTGACTGTGGGTGGGAGTCCCTGATCTCC 562447
QY 134 GGACTTGGCCCCCAGTGTACATGGGCGCAGGGGCACTCCCGTTGCTCTGACTCTGG 193
DB 562448 TTGCTCAGGGCCCGGTGAGACTGTGGGTGAGGAGATCCATGATCTCTGCTCAGG 562507
QY 194 CCCCC 198
DB 562508 GCCCC 562512
```

```
RESULT 7
```

```
US-09-949-016-17361
; Sequence 17361; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
```

/ LOCATION: (1)...(767677)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

Query Match
Best Local Similarity 9.7%; Score 41; DB 4; Length 767677;
51.4%; Pred. No. 0.05;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 14 GGCACCTCCCGTGGTCTCTGAGCTCTGGCCCACTGTATACATGGGCGCACGGGCACTGCC 73
DB 562328 GGGAGTCCCGTATCTCTTCTGCTCAGAGGCCCGTGATGTAAGGTGTAGGAGATCCC 562387
QY 74 GTGGTCTCTGAGCTCTGGCCCGCAGTGTATGATGGCGCACGGGCGCACTCCGTTGGTCTCT 133
DB 562388 GTGATCTCCACACTCAGGGCTCGTGTGACTGTGGGCGTGGGGAGTCCCGTATCTCC 562447
QY 134 GGAATCTGGCCCCCACTGTAACATGGGCGCACGGGCGCACTCCGTTGGTCTCTGAGCTCTGG 193
DB 562448 TTGCTCAGGGGCCCGGAGGAGCTGTGGTGTAGGAAATCCATGATCTCTTGTGCTCAGG 562507
QY 194 CCCCC 198
DB 562508 GCCCC 562512

RESULT 8
US-09-949-016-11786
/ Sequence 11786, Application US/09949016
/ Patent No. 6812339

/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11786
/ LENGTH: 15108
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-11786

Query Match
Best Local Similarity 9.5%; Score 40.2; DB 4; Length 15108;
78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTGGTCTCTGAGCTCTGGCCCACTGTATACATGGGCGCACGGGCA 68
DB 2357 CACGGGCACTCCCGTGGTCTCTGAGCTCTGGCCCGTGTATACATGGGCGTTCCATGAA 2416
QY 69 C 69
DB 2417 C 2417

RESULT 9
US-09-949-016-17205
/ Sequence 17205, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17205
/ LENGTH: 15108
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-17205

Query Match
Best Local Similarity 9.5%; Score 40.2; DB 4; Length 15108;
78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTGGTCTCTGAGCTCTGGCCCGCAGTGTATACATGGGCGCACGGGCA 68
DB 2357 CACGGGCACTCCCGTGGTCTCTGAGCTCTGGCCCGTGTATACATGGGCGTTCCATGAA 2416
QY 69 C 69
DB 2417 C 2417

RESULT 10
US-09-845-583A-1/c
/ Sequence 1, Application US/09845583A
/ Patent No. 6635616
/ GENERAL INFORMATION:
/ APPLICANT: Burgeson, Robert
/ APPLICANT: Brunken, William Joseph
/ APPLICANT: Champilaud, Marie-France
/ APPLICANT: Hunter, Dale
/ TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
/ FILE REFERENCE: 10287-056001
/ CURRENT APPLICATION NUMBER: US/09/845,583A
/ PRIOR FILING DATE: 2001-04-30
/ PRIOR APPLICATION NUMBER: US 60/200,863
/ PRIOR FILING DATE: 2000-05-01
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 11009
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-845-583A-1

Query Match
Best Local Similarity 9.1%; Score 38.4; DB 4; Length 11009;
52.5%; Pred. No. 0.052;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTGGTCTCTGAGCTCTGGCCCGCAGTGTATACATGGGCGCACGGGCA 68
DB 6164 CCGGGGACAGGTGCACTGGCCCGTGTGTGTACAGTGGCTTCGGGACACTGGAG 6105
QY 69 CTCCTGGTCTCTGAGCTCTGGCCCGCAGTGTATACATGGGCGCACGGGCACTCCGTTGG 128
DB 6104 CGCTGGAGCCCTTCTCTGGAGGCCCGCAGTGTAGCAGGGGCGCACTGAGGCACTGGGGT 6045
QY 129 TTCCTGAGCTCTGGCCCGCAGTGTATACATGGGCGCACGGG 168
DB 6044 CCGTGTGCTCTGGCTGTGACATGACAGCCGCTTGGAG 6005

RESULT 11
US-09-949-016-12957
/ Sequence 12957, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:

Dp 13568 GCCACTGTACATCTGGGTGCTCTGGGGACCCGGGGGTTCAAGGATGTTGCCCTGGGT 13627

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

1 1
2 1
3 1
4 1
5 1
6 1
7 1
8 1
9 1
10 1
11 1
12 1
13 1
14 1
15 1
16 1
17 1
18 1
19 1
20 1
21 1
22 1
23 1
24 1
25 1
26 1
27 1
28 1
29 1
30 1
31 1
32 1
33 1
34 1
35 1
36 1
37 1
38 1
39 1
40 1
41 1
42 1
43 1
44 1
45 1
46 1
47 1
48 1
49 1
50 1
51 1
52 1
53 1
54 1
55 1
56 1
57 1
58 1
59 1
60 1
61 1
62 1
63 1
64 1
65 1
66 1
67 1
68 1
69 1
70 1
71 1
72 1
73 1
74 1
75 1
76 1
77 1
78 1
79 1
80 1
81 1
82 1
83 1
84 1
85 1
86 1
87 1
88 1
89 1
90 1
91 1
92 1
93 1
94 1
95 1
96 1
97 1
98 1
99 1
100 1
101 1
102 1
103 1
104 1
105 1
106 1
107 1
108 1
109 1
110 1
111 1
112 1
113 1
114 1
115 1
116 1
117 1
118 1
119 1
120 1
121 1
122 1
123 1
124 1
125 1
126 1
127 1
128 1
129 1
130 1
131 1
132 1
133 1
134 1
135 1
136 1
137 1
138 1
139 1
140 1
141 1
142 1
143 1
144 1
145 1
146 1
147 1
148 1
149 1
150 1
151 1
152 1
153 1
154 1
155 1
156 1
157 1
158 1
159 1
160 1
161 1
162 1
163 1
164 1
165 1
166 1
167 1
168 1
169 1
170 1
171 1
172 1
173 1
174 1
175 1
176 1
177 1
178 1
179 1
180 1
181 1
182 1
183 1
184 1
185 1
186 1
187 1
188 1
189 1
190 1
191 1
192 1
193 1
194 1
195 1
196 1
197 1
198 1
199 1
200 1
201 1
202 1
203 1
204 1
205 1
206 1
207 1
208 1
209 1
210 1
211 1
212 1
213 1
214 1
215 1
216 1
217 1
218 1
219 1
220 1
221 1
222 1
223 1
224 1
225 1
226 1
227 1
228 1
229 1
230 1
231 1
232 1
233 1
234 1
235 1
236 1
237 1
238 1
239 1
240 1
241 1
242 1
243 1
244 1
245 1
246 1
247 1
248 1
249 1
250 1
251 1
252 1
253 1
254 1
255 1
256 1
257 1
258 1
259 1
260 1
261 1
262 1
263 1
264 1
265 1
266 1
267 1
268 1
269 1
270 1
271 1
272 1
273 1
274 1
275 1
276 1
277 1
278 1
279 1
280 1
281 1
282 1
283 1
284 1
285 1
286 1
287 1
288 1
289 1
290 1
291 1
292 1
293 1
294 1
295 1
296 1
297 1
298 1
299 1
300 1
301 1
302 1
303 1
304 1
305 1
306 1
307 1
308 1
309 1
310 1
311 1
312 1
313 1
314 1
315 1
316 1
317 1
318 1
319 1
320 1
321 1
322 1
323 1
324 1
325 1
326 1
327 1
328 1
329 1
330 1
331 1
332 1
333 1
334 1
335 1
336 1
337 1
338 1
339 1
340 1
341 1
342 1
343 1
344 1
345 1
346 1
347 1
348 1
349 1
350 1
351 1
352 1
353 1
354 1
355 1
356 1
357 1
358 1
359 1
360 1
361 1
362 1
363 1
364 1
365 1
366 1
367 1
368 1
369 1
370 1
371 1
372 1
373 1
374 1
375 1
376 1
377 1
378 1
379 1
380 1
381 1
382 1
383 1
384 1
385 1
386 1
387 1
388 1
389 1
390 1
391 1
392 1
393 1
394 1
395 1
396 1
397 1
398 1
399 1
400 1
401 1
402 1
403 1
404 1
405 1
406 1
407 1
408 1
409 1
410 1
411 1
412 1
413 1
414 1
415 1
416 1
417 1
418 1
419 1
420 1
421 1
422 1
423 1
424 1
425 1
426 1
427 1
428 1
429 1
430 1
431 1
432 1
433 1
434 1
435 1
436 1
437 1
438 1
439 1
440 1
441 1
442 1
443 1
444 1
445 1
446 1
447 1
448 1
449 1
450 1
451 1
452 1
453 1
454 1
455 1
456 1
457 1
458 1
459 1
460 1
461 1
462 1
463 1
464 1
465 1
466 1
467 1
468 1
469 1
470 1
471 1
472 1
473 1
474 1
475 1
476 1
477 1
478 1
479 1
480 1
481 1
482 1
483 1
484 1
485 1
486 1
487 1
488 1
489 1
490 1
491 1
492 1
493 1
494 1
495 1
496 1
497 1
498 1
499 1
500 1
501 1
502 1
503 1
504 1
505 1
506 1
507 1
508 1
509 1
510 1
511 1
512 1
513 1
514 1
515 1
516 1
517 1
518 1
519 1
520 1
521 1
522 1
523 1
524 1
525 1
526 1
527 1
528 1
529 1
530 1
531 1
532 1
533 1
534 1
535 1
536 1
537 1
538 1
539 1
540 1
541 1
542 1
543 1
544 1
545 1
546 1
547 1
548 1
549 1
550 1
551 1
552 1
553 1
554 1
555 1
556 1
557 1
558 1
559 1
560 1
561 1
562 1
563 1
564 1
565 1
566 1
567 1
568 1
569 1
570 1
571 1
572 1
573 1
574 1
575 1
576 1
577 1
578 1
579 1
580 1
581 1
582 1
583 1
584 1
585 1
586 1
587 1
588 1
589 1
590 1
591 1
592 1
593 1
594 1
595 1
596 1
597 1
598 1
599 1
600 1
601 1
602 1
603 1
604 1
605 1
606 1
607 1
608 1
609 1
610 1
611 1
612 1
613 1
614 1
615 1
616 1
617 1
618 1
619 1
620 1
621 1
622 1
623 1
624 1
625 1
626 1
627 1
628 1
629 1
630 1
631 1
632 1
633 1
634 1
635 1
636 1
637 1
638 1
639 1
640 1
641 1
642 1
643 1
644 1
645 1
646 1
647 1
648 1
649 1
650 1
651 1
652 1
653 1
654 1
655 1
656 1
657 1
658 1
659 1
660 1
661 1
662 1
663 1
664 1
665 1
666 1
667 1
668 1
669 1
670 1
671 1
672 1
673 1
674 1
675 1
676 1
677 1
678 1
679 1
680 1
681 1
682 1
683 1
684 1
685 1
686 1
687 1
688 1
689 1
690 1
691 1
692 1
693 1
694 1
695 1
696 1
697 1
698 1
699 1
700 1
7
```

